

91473

STIC-Biotech/ChemLib

From: Slobodyansky, Elizabeth  
Sent: Monday, April 14, 2003 11:43 AM  
To: STIC-Biotech/ChemLib  
Subject: 09/527,376

Please search for case 09/527,376:



SEQ ID NOs: 1 and 2 against commercial and interference databases.

Thank you.

*Elizabeth Slobodyansky, PhD*

Primary Examiner

Art Unit 1652  
CM1 10D11  
703-306-3222

**mail box 10D01:**

Mary Jane Ruhl  
Tech. Info. Specialist, STIC  
TC-1600  
CM-1, Room 6A-06  
Phone: 605-1155

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 4/22/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

| Result No. | Score | Match | Query | Length | DB                | ID                 | Description |
|------------|-------|-------|-------|--------|-------------------|--------------------|-------------|
|            |       |       |       |        |                   |                    |             |
| 1          | 511   | 52.1  | 198   | 4      | US-09-045-973-5   | Sequence 5, Appli  |             |
| 2          | 275   | 28.0  | 226   | 4      | US-08-045-973-8   | Sequence 8, Appli  |             |
| 3          | 252   | 25.7  | 367   | 2      | US-09-530-290-24  | Sequence 24, Appli |             |
| 4          | 252   | 25.7  | 367   | 2      | US-08-990-379-6   | Sequence 6, Appli  |             |
| 5          | 240   | 24.5  | 394   | 2      | US-08-530-290-23  | Sequence 23, Appli |             |
| 6          | 236   | 24.1  | 393   | 2      | US-08-990-379-4   | Sequence 4, Appli  |             |
| 7          | 236   | 24.1  | 395   | 2      | US-08-990-379-5   | Sequence 5, Appli  |             |
| 8          | 222   | 22.6  | 176   | 4      | US-09-704-139-2   | Sequence 2, Appli  |             |
| 9          | 221   | 22.5  | 397   | 2      | US-08-990-379-8   | Sequence 8, Appli  |             |
| 10         | 220   | 22.4  | 173   | 4      | US-09-704-139-4   | Sequence 4, Appli  |             |
| 11         | 216   | 22.0  | 117   | 1      | US-07-988-273-4   | Sequence 4, Appli  |             |
| 12         | 216   | 22.0  | 117   | 5      | PCR-US93-12019-4  | Sequence 4, Appli  |             |
| 13         | 213.5 | 21.8  | 172   | 4      | US-09-704-139-5   | Sequence 5, Appli  |             |
| 14         | 208   | 21.2  | 313   | 2      | US-08-990-379-7   | Sequence 7, Appli  |             |
| 15         | 208   | 21.2  | 314   | 4      | US-09-164-193-22  | Sequence 22, Appli |             |
| 16         | 208   | 21.2  | 314   | 4      | US-09-221-448A-22 | Sequence 22, Appli |             |
| 17         | 194.5 | 19.8  | 198   | 4      | US-09-163-833-2   | Sequence 2, Appli  |             |
| 18         | 179.5 | 18.3  | 207   | 4      | US-09-013-881-4   | Sequence 4, Appli  |             |
| 19         | 171   | 17.4  | 263   | 4      | US-09-164-193-5   | Sequence 5, Appli  |             |
| 20         | 171   | 17.4  | 263   | 4      | US-09-221-448A-5  | Sequence 5, Appli  |             |
| 21         | 162.5 | 16.6  | 185   | 1      | US-07-988-273-2   | Sequence 2, Appli  |             |
| 22         | 162.5 | 16.6  | 185   | 3      | US-08-848-810-25  | Sequence 25, Appli |             |
| 23         | 162.5 | 16.6  | 185   | 4      | US-09-164-193-21  | Sequence 21, Appli |             |
| 24         | 162.5 | 16.6  | 185   | 4      | US-09-221-448A-21 | Sequence 21, Appli |             |
| 25         | 162.5 | 16.6  | 185   | 5      | PCR-US93-12019-2  | Sequence 2, Appli  |             |
| 26         | 146   | 14.9  | 72    | 2      | US-08-530-290-19  | Sequence 19, Appli |             |
| 27         | 141   | 14.4  | 353   | 4      | US-09-013-881-3   | Sequence 3, Appli  |             |





```
Db 163 SCTPLY-----DOGGVEILPFLYLGSAVHASRKM LMDALGITALINVSANCNPFHEGY 218
QY 65 QYMQVPVADSPNRLCDFEPIADHITHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHA 124
Db 219 QYKSIPEVDNHKADISSWNEAIDFDSIKNAGGRVFVHCQAGISRSATICLAYLMRTNR 278
QY 125 MSLLDAHTWTKSCRPIIRPNSGFEQLIHFEOLFQKNTVHMVSSPVGMIPD 176
Db 279 VKLDEAFEFVKQRSSIISPNFSPMGQLLOFESQVLAPHCSAEAGSPAMAVLD 330

RESULT 4
US-08-990-379-6
; Sequence 6, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Phillip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; EARLIER FILING DATE: 1997-12-15
; EARLIER FILING DATE: 1996-06-14
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Mus sp.
US-08-990-379-6

Query Match 25.7%; Score 252; DB 2; Length 367;
Best Local Similarity 34.3%; Pred. No. 1.4e-22;
Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps 1;

QY 5 SCAPPVQFRQPSVSGLSQITKSLYISNGVAANKMLSSNQITPMVINSVVNTLYEDI 64
Db 163 SCTPLY-----DOGGVEILPFLYLGSAVHASRKM LMDALGITALINVSANCNPFHEGY 218
QY 65 QYMQVPVADSPNRLCDFEPIADHITHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHA 124
Db 219 QYKSIPEVDNHKADISSWNEAIDFDSIKNAGGRVFVHCQAGISRSATICLAYLMRTNR 278
QY 125 MSLLDAHTWTKSCRPIIRPNSGFEQLIHFEOLFQKNTVHMVSSPVGMIPD 176
Db 279 VKLDEAFEFVKQRSSIISPNFSPMGQLLOFESQVLAPHCSAEAGSPAMAVLD 330

RESULT 5
US-08-530-290-23
; Sequence 23, Application US/08530290
; Patent No. 5958721
; GENERAL INFORMATION:
; APPLICANT: Marshall, Christopher John
; APPLICANT: Ashworth, Alan
; APPLICANT: Hughes, David Anthony
; TITLE OF INVENTION: Methods for Screening of Substances for
; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,290
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: WO PCT/GB94/00694
; FILING DATE: 31-MAR-1994
; PRIOR APPLICATION NUMBER: GB 9402573.1
; FILING DATE: 10-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9307250.2
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 084611-000000005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-530-290-23

Query Match 24.5%; Score 240; DB 2; Length 394;
Best Local Similarity 33.9%; Pred. No. 4.8e-21;
Matches 57; Conservative 32; Mismatches 75; Indels 4; Gaps 1;

QY 5 SCAPPVQFRQPSVSGLSQITKSLYISNGVAANKMLSSNQITPMVINSVVNTLYEDI 64
Db 185 SCTPLHDQE---GPVEILPFLYLGSAVHAARRMDLALGITALLNVSSDCPNFHEGY 240
QY 65 QYMQVPVADSPNRLCDFEPIADHITHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHA 124
Db 241 QYKCIPEVDNHKADISSWFEALIEYIDAVKDCRGRVLVHCQAGISRSATICLAYLMKKR 300
QY 125 MSLLDAHTWTKSCRPIIRPNSGFEQLIHFEOLFQKNTVHMVSSPVG 172
Db 301 VRLEAEFEVKQRSSIISPNFSPMGQLLOFESQVLATSCAAEAASPSG 348

RESULT 6
US-08-990-379-4
; Sequence 4, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Phillip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-08-990-379-4

Query Match 24.1%; Score 236; DB 2; Length 393;
```

```

Best Local Similarity 33.9%; Pred. No. 1.5e-20;
Matches 57; Conservative 30; Mismatches 77; Indels 4; Gaps 1;

QY      5 SCAPVQFRQPSVSGLSQITKSLYSINCVAAANNKMLSSNOITWTVINSVEVWNTLYEDI 64
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     184 SCGTPLH-----DQGPGVEILPFLVLGSAYHAARDMDLDALGITALLNVSSDCPNHFEGHY 239
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      65 QYMOPVPADSPNSRLCDDFPDIADIIHSVENMKOGRTHLLHCAAGVSRSAAALCLAYLMKYHA 124
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     240 QYKICIPVEDNHKAIDISWFMEAEIYDAVKDCRGLVHCQAGISRATICTLAYLMKKR 299
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      125 MSLLDAHTWTKSCRPIIRNPNSGFWEQLIHIEFQFGKNVTVMHVSSPVG 172
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     300 VRLEAEFEFKQRSIIISPNSFFMQQLIQFESQVLTTSCAAEAASPSG 347
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 7
US-08-990-379-5
; Sequence 5, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Philip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase CDNAs and
; TITLE OF INVENTION: Their Biologically Active Expression Products
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-08-990-379-5

Query Match          24.1%; Score 236; DB 2; Length 395;
Best Local Similarity 33.9%; Pred. No. 1.5e-20;
Matches 57; Conservative 30; Mismatches 77; Indels 4; Gaps 1;

QY      5 SCAPVQFRQPSVSGLSQITKSLYSINCVAAANNKMLSSNOITWTVINSVEVWNTLYEDI 64
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     186 SCGTPLH-----DQGPGVEILPFLVLGSAYHAARDMDLDALGITALLNVSSDCPNHFEGHY 241
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      65 QYMOPVPADSPNSRLCDDFPDIADIIHSVENMKOGRTHLLHCAAGVSRSAAALCLAYLMKYHA 124
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     242 QYKICIPVEDNHKAIDISWFMEAEIYDAVKDCRGLVHCQAGISRATICTLAYLMKKR 301
       || | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      125 MSLLDAHTWTKSCRPIIRNPNSGFWEQLIHIEFQFGKNVTVMHVSSPVG 172
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     302 VRLEAEFEFKQRSIIISPNSFFMQQLIQFESQVLTTSCAAEAASPSG 349
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 8
US-09-704-139-2
; Sequence 2, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES T
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 176

```

FILE REFERENCE: 10448-018001  
CURRENT APPLICATION NUMBER: US/09/704,139  
CURRENT FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: US 60/185,772  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 4  
LENGTH: 173  
TYPE: PRT  
ORGANISM: Artificial/Unknown  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)..(173)  
OTHER INFORMATION: consensus sequence  
US-09-704-139-4

Query Match 22.4%; Score 220; DB 4; Length 173;  
Best Local Similarity 34.9%; Pred. No. 4e-19;  
Matches 60; Conservative 21; Mismatches 55; Indels 36; Gaps 5;

QY 19 GLSQTSLYISNGVAAN--NKLMLSSNOITMVINVSVEVNTLYED----- 63  
Db 1 GFSEILPHYLGSYSTASEANLALKLKGITHVINTEVPNPFELDKKNDRHYNAYIS 60  
QY 64 ----IQMQVP-----VADSPNSRLCDFDPIADHIHVSVMKQGRITLLHCAAGVSR 110  
Db 61 KNSGFTYLIQPNVDHIIYHIAWNETKISKYFDEAVDFIDAROKGKVLVHCQAGISR 120  
QY 111 SAAALCLAYLMKYHMSLLDA-----HPWTKSCRPIIRNPSGFWQLIHYE 155  
Db 121 SATLIATILMKTRNLNSLAYDFVYVYHIKERRC-PIISNFGFLRLIYE 171

RESULT 11  
US-07-988-273-4  
Sequence 4, Application US/07988273  
Patent No. 5512434  
GENERAL INFORMATION:  
APPLICANT: AARONSON, Stuart A.  
APPLICANT: BOTTARO, Donald P.  
APPLICANT: ISHIBASHI, Toshio  
APPLICANT: MIKI, Toru  
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN  
TITLE OF INVENTION: PHOSPHATASE  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,273  
FILING DATE: 19921214  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40399/182 NIHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 117 amino acids  
TYPE: AMINO ACID

TOPOLOGY: linear  
US-07-988-273-4

Query Match 22.0%; Score 216; DB 1; Length 117;  
Best Local Similarity 39.7%; Pred. No. 6.8e-19;  
Matches 46; Conservative 20; Mismatches 50; Indels 0; Gaps 0;

QY 40 MLSSNOITMVINVSVEVNTLYEDIQMQVPVADSPNSRLCDFDPIADHIHVSVMKQGR 99  
Db 1 MLDALGITALINVSANCPNHFEGHYQKSIPIVEDNHNKADISSWFNEAIDFIDSIKDAGGR 60  
QY 100 TLLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPIIRNPSGFWQLIHYE 155  
Db 61 VFVHCQAGISRATICLAYLMRTNRVKLDEAFEFVKORRSIIISPNFSFMGQLLOQE 116

RESULT 12  
PCT-US93-12019-4  
Sequence 4, Application PC/TUS9312019  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN  
TITLE OF INVENTION: PHOSPHATASE  
NUMBER OF SEQUENCES: 7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/12019  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,273  
FILING DATE: 14-DEC-1992  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 117 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US93-12019-4

Query Match 22.0%; Score 216; DB 5; Length 117;  
Best Local Similarity 39.7%; Pred. No. 6.8e-19;  
Matches 46; Conservative 20; Mismatches 50; Indels 0; Gaps 0;

QY 40 MLSSNOITMVINVSVEVNTLYEDIQMQVPVADSPNSRLCDFDPIADHIHVSVMKQGR 99  
Db 1 MLDALGITALINVSANCPNHFEGHYQKSIPIVEDNHNKADISSWFNEAIDFIDSIKDAGGR 60  
QY 100 TLLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPIIRNPSGFWQLIHYE 155  
Db 61 VFVHCQAGISRATICLAYLMRTNRVKLDEAFEFVKORRSIIISPNFSFMGQLLOQE 116

RESULT 13  
US-09-704-139-5  
Sequence 5, Application US/09704139  
Patent No. 6420153  
GENERAL INFORMATION:  
APPLICANT: Kapeller, Rosanna  
TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREOF  
FILE REFERENCE: 10448-018001  
CURRENT APPLICATION NUMBER: US/09/704,139  
CURRENT FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: US 60/185,772  
PRIOR FILING DATE: 2000-02-29  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 5  
LENGTH: 172  
TYPE: PRT  
ORGANISM: Artificial/Unknown

; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)..(172)  
; OTHER INFORMATION: consensus sequence  
US-09-704-139-5

Query Match 21.8%; Score 213.5; DB 4; Length 172;  
Best Local Similarity 34.5%; Pred. No. 2.5e-18;  
Matches 59; Conservative 21; Mismatches 56; Indels 35; Gaps 5;  
QY 19 GLSQITKSLYISNGVAAN--NKLMLSSNQITWTVNVSVEVNTLY----- 61  
DB 1 GPSELPHLYGSYSDASEANLALKLIGTHVINVTVEENFELKKKNDRYTNEYIS 60  
QY 62 --EDIQMVDP-----VADSPNSRLCDFDPIADHIHSVEMKOGRTLLHCAAGVSR 111  
DB 61 KSGFTYLIQIPNVDDIYHIAWNTETKISKYLEEAEVIEDAEKKGKVLVHCAAGVSR 120  
QY 112 AALCLAYLMKYHAMSLDLA-----HTWTKSCRPIIRPNSGFWQELIHYE 155  
DB 121 ATLVIAYLMKTRNLSLRDAYDEVYVYHIKERRC--PIISPNFGFLRLIEYE 170

## RESULT 14

US-08-990-379-7

; Sequence 7, Application US/08990379

; Patent No. 5998188

; GENERAL INFORMATION:

; APPLICANT: Stork, Philip J

; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and

; FILE REFERENCE: 4104-000322USA

; CURRENT APPLICATION NUMBER: US/08/990,379

; EARLIER FILING DATE: 1997-12-15

; EARLIER FILING DATE: 1996-06-14

; EARLIER FILING DATE: 1995-08-16

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 313

; TYPE: PRT

; ORGANISM: Mus sp.

US-08-990-379-7

Query Match 21.2%; Score 208; DB 2; Length 313;  
Best Local Similarity 33.1%; Pred. No. 3e-17;  
Matches 55; Conservative 26; Mismatches 75; Indels 10; Gaps 1;

QY 3 APSCAFPVQFRPQSVS-----GLSQITKSLYISNGVAANNNKMLSSNQITWTVN 52  
DB 145 APAQALPPAGAENSNDPRVPIYDQGGPVEILPYLYLGSCNHSDDLQGLQACGITA 204  
QY 53 SVEVNTLYEDIQYMQVPVADSPNSRLCDFDPIADHIHSVEMKOGRTLLHCAAGVSR 112  
DB 205 SASCPNHFEGFLFHYKSIPTVEDNQMVETSAWFEAISFTDSVKNSGGRVLVHCAAG 264  
QY 113 ALCLAYLMKYHAMSLDLAHTWTKSCRPIIRPNSGFWQELIHYEFQL 158  
DB 265 TICLAYLIQSHRVLDEAFDFVKORRGVISPNSFGMLQLQLETVQV 310

## RESULT 15

US-09-164-193-22

; Sequence 22, Application US/09164193C

; Patent No. 6258582

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan L.

; TITLE OF INVENTION: NOVEL CSATPP NUCLEIC ACID MOLECULES AND USES THEREFOR (as amended

; FILE REFERENCE: MNI-051

; CURRENT APPLICATION NUMBER: US/09/164,193C

; CURRENT FILING DATE: 1998-09-30  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-164-193-22

Query Match 21.2%; Score 208; DB 4; Length 314;  
Best Local Similarity 33.1%; Pred. No. 3e-17;  
Matches 55; Conservative 26; Mismatches 75; Indels 10; Gaps 1;  
QY 3 APSCAFPVQFRPQSVS-----GLSQITKSLYISNGVAANNNKMLSSNQITWTVN 52  
DB 146 APAQALPPAGAENSNDPRVPIYDQGGPVEILPYLYLGSCNHSDDLQGLQACGITA 205  
QY 53 SVEVNTLYEDIQYMQVPVADSPNSRLCDFDPIADHIHSVEMKOGRTLLHCAAGVSR 112  
DB 206 SASCPNHFEGFLFHYKSIPTVEDNQMVETSAWFEAISFTDSVKNSGGRVLVHCAAG 265  
QY 113 ALCLAYLMKYHAMSLDLAHTWTKSCRPIIRPNSGFWQELIHYEFQL 158  
DB 266 TICLAYLIQSHRVLDEAFDFVKORRGVISPNSFGMLQLQLETVQV 311

Search completed: April 21, 2003, 18:35:37

Job time : 23 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 18:23:34 ; Search time 44 Seconds  
(without alignments)  
880.383 Million cell m

Title: US-09-527-376-2

Perfect score: 981

Sequence: 1 MTAPSCAFPVQFRQPSVSL.....SPVGMIPDIYEKEVRLMIPL 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

For \_\_\_\_\_

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

- ```

1:  sp.archaea:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp.human:*
5:  sp.invertebrate:*
6:  sp.mammal:*
7:  sp.mnc:*
8:  sp.organelle:*
9:  sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.invertebrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query % |     | Length | DB     | ID                  | Description |
|------------|-------|---------|-----|--------|--------|---------------------|-------------|
|            |       | Match   |     |        |        |                     |             |
| 1          | 818   | 83.4    | 188 | 11     | Q8VE01 | mus musculus        |             |
| 2          | 780   | 79.5    | 151 | 4      | Q8BTC8 | Q9btg8 homo sapien  |             |
| 3          | 733.5 | 74.8    | 189 | 11     | Q9D9D8 | Q949d8 mus musculus |             |
| 4          | 700   | 71.4    | 190 | 4      | Q9H596 | Q9h596 homo sapien  |             |
| 5          | 661   | 67.4    | 190 | 4      | Q9I346 | Q9y346 homo sapien  |             |
| 6          | 338   | 34.5    | 272 | 5      | Q9XVB7 | Q9xv87 caenorhabdi  |             |
| 7          | 304   | 31.0    | 212 | 5      | Q9VAB0 | Q9vab0 drosophila   |             |
| 8          | 275   | 28.0    | 226 | 5      | Q93592 | Q93592 caenorhabdi  |             |
| 9          | 253   | 25.8    | 353 | 13     | Q42253 | Q42253 gallus gall  |             |
| 10         | 252   | 25.7    | 367 | 11     | Q64193 | Q64193 rattus sp.   |             |
| 11         | 252   | 25.7    | 367 | 11     | Q63683 | Q63683 rattus norv  |             |
| 12         | 252   | 25.7    | 369 | 13     | Q91790 | Q91790 xenopus lae  |             |
| 13         | 249   | 25.4    | 369 | 13     | Q90W58 | Q90w58 xenopus lae  |             |
| 14         | 245   | 25.0    | 476 | 5      | Q46132 | Q46122 drosophila   |             |
| 15         | 240   | 24.5    | 411 | 4      | Q13649 | Q13649 homo sapien  |             |
| 16         | 240   | 24.5    | 476 | 5      | Q9VHV8 | Q9vhv8 drosophila   |             |

|    |       |      |     |    |        |         |              |
|----|-------|------|-----|----|--------|---------|--------------|
| 17 | 238   | 24.3 | 303 | 4  | Q9NSW1 | Q9nsw1  | homo sapien  |
| 18 | 233   | 23.8 | 606 | 5  | Q8ST18 | Q8st18  | caenorhabdi  |
| 19 | 233   | 23.8 | 657 | 5  | Q8ST19 | Q8st19  | caenorhabdi  |
| 20 | 231.5 | 23.6 | 186 | 5  | Q19388 | Q19388  | caenorhabdi  |
| 21 | 223.5 | 22.8 | 365 | 5  | Q44128 | Q44128  | caenorhabdi  |
| 22 | 220   | 22.4 | 198 | 10 | Q9LUG6 | Q9lug6  | arabidopsis  |
| 23 | 222   | 22.4 | 198 | 10 | Q9R37  | Q9r37   | arabidopsis  |
| 24 | 217   | 22.1 | 217 | 4  | Q8WTR2 | Q8wtr2  | homo sapien  |
| 25 | 216   | 22.0 | 227 | 5  | Q95XK5 | Q95xk5  | caenorhabdi  |
| 26 | 214   | 21.8 | 184 | 11 | Q99N11 | Q99n11  | mus musculus |
| 27 | 208.5 | 21.3 | 660 | 11 | Q920R2 | Q920r2  | mus musculus |
| 28 | 208   | 21.3 | 677 | 11 | Q95MG6 | Q95mg6  | mus musculus |
| 29 | 207   | 21.1 | 184 | 4  | Q9NRW4 | Q9nrw4  | homo sapien  |
| 30 | 205.5 | 20.9 | 665 | 4  | Q96N49 | Q96n49  | homo sapien  |
| 31 | 205   | 20.9 | 206 | 5  | Q8SXF8 | Q8sxf8  | drosophila   |
| 32 | 204   | 20.8 | 167 | 10 | Q9M8K7 | Q9m8k7  | arabidopsis  |
| 33 | 203   | 20.7 | 173 | 11 | Q63758 | Q63758  | rattus sp.   |
| 34 | 201   | 20.5 | 220 | 11 | Q95N12 | Q95n12  | mus musculus |
| 35 | 200   | 20.4 | 162 | 11 | Q9CRR3 | Q9crr3  | mus musculus |
| 36 | 200   | 20.4 | 177 | 11 | Q9CSL5 | Q9csl5  | mus musculus |
| 37 | 200   | 20.4 | 483 | 11 | Q9CZY9 | Q9czy9  | mus musculus |
| 38 | 200   | 20.4 | 483 | 11 | Q8CRL3 | Q8crl3  | mus musculus |
| 39 | 198.5 | 20.2 | 211 | 4  | Q9BV47 | Q9bv47  | homo sapien  |
| 40 | 198   | 20.2 | 220 | 11 | Q9P6P6 | Q9dp6p6 | mus musculus |
| 41 | 198   | 20.2 | 483 | 11 | Q9ESS0 | Q9ess0  | mus musculus |
| 42 | 196   | 20.0 | 348 | 5  | Q9VVM4 | Q9vvw4  | drosophila   |
| 43 | 196   | 20.0 | 411 | 5  | Q95SV1 | Q95sv1  | drosophila   |
| 44 | 195   | 19.9 | 188 | 4  | Q96367 | Q96j67  | homo sapien  |
| 45 | 194.5 | 19.8 | 784 | 10 | Q9C5S1 | Q9c5s1  | arabidopsis  |

## ALIGNMENTS

## RESULT 1

|    |                                                                   |                                         |      |         |
|----|-------------------------------------------------------------------|-----------------------------------------|------|---------|
| AC | Q8VE01                                                            | PRELIMINARY;                            | PRT; | 188 AA. |
| ID | Q8VE01                                                            |                                         |      |         |
| IC | Q8VE01                                                            |                                         |      |         |
| DD | 01-MAR-2002                                                       | (TREMBLrel. 20, Created)                |      |         |
| DT | 01-MAR-2002                                                       | (TREMBLrel. 20, Last sequence update)   |      |         |
| DT | 01-JUN-2002                                                       | (TREMBLrel. 21, Last annotation update) |      |         |
| DE | Hypothetical 21.1 kDa protein.                                    |                                         |      |         |
| OS | Mus musculus (Mouse).                                             |                                         |      |         |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;           |                                         |      |         |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus |                                         |      |         |
| OX | NCBI_TaxID=10090;                                                 |                                         |      |         |
| RP | [1]                                                               |                                         |      |         |
| RN | SEQUENCE FROM N.A.                                                |                                         |      |         |
| RA | Strausberg R.;                                                    |                                         |      |         |
| RL | Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.          |                                         |      |         |
| RR | EMBL: BC020036; AAH20036.1; -                                     |                                         |      |         |
| DR | InterPro: IPR000340; DS_phosphatase.                              |                                         |      |         |
| DR | InterPro: IPR000387; Tyr_phosphatase.                             |                                         |      |         |
| DR | Pfam: PF00782; DSPC; 1.                                           |                                         |      |         |
| DR | SMART: SM00195; DSPC; 1.                                          |                                         |      |         |
| DR | PROSITE: PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.                   |                                         |      |         |
| DR | PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.                           |                                         |      |         |
| DR | PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.                        |                                         |      |         |
| DR | Hypothetical protein.                                             |                                         |      |         |
| SW | SEQUENCE 188 AA; 21118 MW; 86CB43E390DC7DE3 CRC64;                |                                         |      |         |

Query Match 83.4%: Score 818: DB 11: Length 188:

```
Qy 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFGKNTVHMVSSPVGMPIDPIYK 180
Db 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFGKNTVHMVSSPVGMPIDPIYK 180
Qy 181 EVRLMPL 188
Db 181 ETRLMPL 188

RESULT 2
Q9BTG8 ID Q9BTG8 PRELIMINARY; PRT; 151 AA.
AC Q9BTG8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 17.1 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004110; AA04110.1; -
DR HSP; P51452; 1VHR.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 151 AA; 17111 MW; 43B60648E7023FDF CRC64;

Query Match 79.5%; Score 780; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.3e-74;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 LSSNOITWVINVSVVNTLYEDIQYMOVPVADSPNSRLCDFDPIADHIHSEVMKQGR 100
Db 4 LSSNOITWVINVSVVNTLYEDIQYMOVPVADSPNSRLCDFDPIADHIHSEVMKQGR 63
Qy 101 LLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFG 160
Db 64 LLHCAAGVSRSAALCLAYLMKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFG 123
Qy 161 KNTVHMVSSPVGMPIDPIYKEVRLMPL 188
Db 124 KNTVHMVSSPVGMPIDPIYKEVRLMPL 151

RESULT 3
Q9D9D8 ID Q9D9D8 PRELIMINARY; PRT; 189 AA.
AC Q9D9D8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1700094E07R1k protein.
GN 1700094E07R1k.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;

Qy 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFGKNTVHMVSSPVGMPIDPIYK 180
Db 121 KYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFGKNTVHMVSSPVGMPIDPIYK 180
Qy 181 EVRLMPL 188
Db 181 ETRLMPL 188

MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007061; BAR24847.1; -
DR HSP; P51452; 1VHR.
DR MGD; MGI:1920797; 1700094E07R1k.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 189 AA; 21508 MW; 8EE2F7151BB5CBFA CRC64;

Query Match 74.8%; Score 733.5; DB 11; Length 189;
Best Local Similarity 73.7%; Pred. No. 1.4e-69;
Matches 137; Conservative 21; Mismatches 27; Indels 1; Gaps 1;

Qy 1 MTAPSCAFFVQ-FRPPSVGLSJOITKSLYISNGVAANNKMLSSNOITWVINVSVVNT 59
Db 1 MTAPSCAFFVQ-FRPPSVGLSJOITKSLYISNGVAANNKMLSSNOITWVINVSVVNT 60
Qy 60 LTEDIQYMOVPVADSPNSRLCDFDPIADHIHSEVMKQGRLLHCAAGVSRSAALCLAYL 119
Db 61 PFEDIQYVQVPVADSPNSLYDFDPIADHIHSEVMKQGRLLHCAAGVSRSAALCLAYL 120
Qy 120 MKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFGKNTVHMVSSPVGMPIDPIY 179
Db 121 MKYHMSLLDAHTWTKSCRPPIRPNNGFWQLIHYEFQFGKNTVHMVSSPVGMPIDPIY 180
Qy 180 KEVRLM 185
Db 181 KEVRLM 186

RESULT 4
Q9H596 ID Q9H596 PRELIMINARY; PRT; 190 AA.
AC Q9H596
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BA386N14.1 (novel protein similar to a dual specificity
DE phosphatase).
GN BA386N14.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
```

```

DR EMBL; AL133545; CAC10195.1; -
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 190 AA; 21529 MW; 3E52BA31A4944EE3 CRC64;

Query Match 71.4%; Score 700; DB 4; Length 190;
Best Local Similarity 68.9%; Pred. No. 5.1e-66;
Matches 131; Conservative 32; Mismatches 25; Indels 2; Gaps 1;

QY 1 MTAPSCAFPVQ--FRQPSVSGLSQITKSLYISNGVAANKMLSSNQITVMVNSVEVYN 58
Db 1 MTASASSFSFGVQVQPSIYSFQITRSLFLSNGVAANDKLLSSNRITAIYNASVGSQ 60

QY 59 TLYEDIQYQVVPVADPSNRLCDFDPIADHHSVEMKQRTLLHCAAGVSRSAALCLAY 118
Db 61 VFEFGIYIKVPVTDARDSRLYDFDPIADLTIHTIDMROGRTLLHCAAGVSRSAICLAY 120

QY 119 LMKYHMSLLDAHTWTKSRPIIRPNSGFWEQLIHYEFOLFGKNTVHMVSSPVGMIPDIY 178
Db 121 LMKYHMSLLDAHTWTKSRPIIRPNSGFWEQLIHYEFKLFNNNTVYRMINSPVGNIPDIY 180

QY 179 EKEVRLMIPL 188
Db 181 EKDLRTMISM 190

RESULT 5
QY346
ID QY346 PRELIMINARY; PRT; 190 AA.
AC QY346;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DE Hypothetical 21.4 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Persson A.E., Sivertson M., Lundeberg J., Uhlen M.;
RT "EU-IMAGE: Full-insert length sequencing of human cDNA clones.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF143321; AAD33910.1; -.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
SQ SEQUENCE 190 AA; 21402 MW; 509A9BB850608712 CRC64;

Query Match 67.4%; Score 661; DB 4; Length 190;
Best Local Similarity 65.8%; Pred. No. 6.8e-62;
Matches 125; Conservative 32; Mismatches 31; Indels 2; Gaps 1;

QY 1 MTAPSCAFPVQ--FRQPSVSGLSQITKSLYISNGVAANKMLSSNQITVMVNSVEVYN 58
Db 1 MTASASSFSFGVQVQPSIYSFQITRSLFLSNGVAANDKLLSSNRITAIYNASVGSQ 60

QY 59 TLYEDIQYQVVPVADPSNRLCDFDPIADHHSVEMKQRTLLHCAAGVSRSAALCLAY 118
Db 61 RILRGLQYIKVPVTDARDSRLYDFDPIADLTIHTIDMROGRTLLHCAAGVSRSAICLAY 120

QY 119 LMKYHMSLLDAHTWTKSRPIIRPNSGFWEQLIHYEFOLFGKNTVHMVSSPVGMIPDIY 178

```

```

Db 121 LMKYHMSLLDAHTWTKSRPIIRPNSGFWEQLIHYEFKLFNNNTVYRMINSPVGNIPDIY 180
QY 179 EKEVRLMIPL 188
Db 181 EKDLRTMISM 190

RESULT 6
QY347
ID QY347 PRELIMINARY; PRT; 272 AA.
AC QY347;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE CO4F12.8 protein.
GN CO4F12.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81461; CAB03837.1; -.
DR HSP; Q16828; MKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00012; PTPC_DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS00056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS00054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ SEQUENCE 272 AA; 30609 MW; 79DD565BBC3F09F0 CRC64;

Query Match 34.5%; Score 338; DB 5; Length 272;
Best Local Similarity 40.7%; Pred. No. 1.6e-27;
Matches 72; Conservative 40; Mismatches 59; Indels 6; Gaps 4;

QY 13 RQPSVSG-LSQITKSLYISN-GVAANKMLSSNQITVMVNSVEVNTLYEDIQYMOVP 70
Db 3 RNWGLGQSEINDHLYSGAGVLAPDK--IKQRKINMIVNATTEPSTYMOGVDTKIR 60

QY 71 VADSPNSRLCDFDPIADHHSVEMKQRTLLHCAAGVSRSAALCLAYLMKYHMSLLDA 130
Db 61 IEDHPYARLNEHFDVVADKIRNKEGGKTLVHCMAAGVSRASLVMILVKHEHMTLROA 120

QY 131 HTWTKSCRPIIRPNSGFWEQLIHYEFOLFGKNTVHMVSSPV--GMIPDIYKEVRUM 185
Db 121 YHYVKAARPIIRPNVGFQWQMDYKRLGRGTASVKMVQTPCEDMPIPDVYADDIRM 177

RESULT 7
QY348
ID QY348 PRELIMINARY; PRT; 212 AA.
AC QY348;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE CG15528 protein.
GN CG15528.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```





```

OX NCBI_TaxID=10118;
RN [1]
SEQUENCE FROM N.A.
RA MEDLINE=96163253; PubMed=8587253;
RX Feng L., Xia Y., Seiffert D., Wilson C.B.;
RT "Oxidative stress-inducible protein tyrosine phosphatase in
RL glomerulonephritis.";
RR Kldney Int. 48:1920-1928(1995).
DR EMBL: S81478; AAB36123.1; -.
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolase.
SQ
SEQUENCE 367 AA; 39580 MW; EEE34CB4E04995A0 CRC64;

Query Match 25.7%; Score 252; DB 11; Length 367;
Best Local Similarity 34.3%; Pred. No. 3.le-18;
Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps

QY 5 SCAPPVQFRPSYGLSQLTKSLYSINGVAANKMLSSNQLITWLVNSVEVVNTLYEDI 64
DB 163 SCTSPY-----DGGGPVEILSFLYGSAHSRKDLMDALGITALLINVSANCNPHEGHY 124
QY 65 QYMQVPVADSPNSRLCDFDPDIADHLSHVEMKGRILLHCAAGVSRSAAALCLAYLMKYHA 124
DB 219 QKSIPVEDNHKAIDISWFNEAIDFDSIKDAGRFEVHCQACIGSRASATICLAYLMRTNR 278
QY 125 MSLLDADTWTSCRPRIPIRPNSGFWGEOHLHYEFOLFGKNTHVMVSSPVGMIPD 176
DB 279 VKLDEAFEFVKQRSIISPNSFPNGQLLOPFESQVLAPHCSAEAGSPAMAVLD 330

RESULT 11
Q63683
ID AC Q63683 PRELIMINARY; PRT; 367 AA.
AC Q63683;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Protein tyrosine phosphatase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
RX MEDLINE=99396926; PubMed=10467595;
RA Qian Z., Gilbert M., Kandel E.R.;
RT "Temporal and spatial regulation of the expression of BAD2, a MAP
RT kinase phosphatase, during seizure, kindling, and long-term
RL potentiation.";
RR Learn. Mem. 1:180-188(1994).
DR EMBL: U02553; AAA03432.1; -.
DR HSSP: Q16828; IMKP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR001763; Rhodanese-like.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR Pfam: PF00581; Rhodanese; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 17:54:24 ; Search time 43 Seconds  
(without alignments)  
582.584 Million cell updates/sec

Title: US-09-527-376-2

Perfect score: 981

Sequence: 1 MTPASCAFFPQFRQPSVSL.....SPVGMIPDIYEKEVRLMIPL 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_101002:\*
- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
  - 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
  - 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
  - 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
  - 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
  - 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
  - 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
  - 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
  - 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
  - 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
  - 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
  - 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
  - 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
  - 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
  - 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
  - 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
  - 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
  - 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
  - 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
  - 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
  - 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
  - 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
  - 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID          | Description          |
|------------|-------|-------------|--------|-------------|----------------------|
| 1          | 981   | 100.0       | 188    | 21 AAB19008 | A human dual-speci   |
| 2          | 981   | 100.0       | 188    | 22 AAM39323 | Human polypeptide    |
| 3          | 981   | 100.0       | 188    | 22 AAB85360 | Human phosphatase    |
| 4          | 981   | 100.0       | 188    | 22 AAB73221 | Human phosphatase    |
| 5          | 981   | 100.0       | 192    | 22 AAU22098 | Human cardiovascular |
| 6          | 981   | 100.0       | 192    | 22 AAM41109 | Human polypeptide    |
| 7          | 981   | 100.0       | 192    | 22 AAB21820 | Novel human neopla   |
| 8          | 921   | 93.9        | 177    | 22 ABB29053 | Peptide #1704 enco   |
| 9          | 921   | 93.9        | 177    | 22 ABB34215 | Peptide #1721 enco   |
| 10         | 921   | 93.9        | 177    | 22 AAM55008 | Human brain expres   |

|    |       |      |     |             |                    |
|----|-------|------|-----|-------------|--------------------|
| 11 | 921   | 93.9 | 177 | 22 AAM67394 | Human bone marrow  |
| 12 | 921   | 93.9 | 177 | 22 AAM15224 | Peptide #1658 enco |
| 13 | 921   | 93.9 | 177 | 22 AAM27685 | Peptide #1722 enco |
| 14 | 921   | 93.9 | 177 | 22 AAM02967 | Peptide #1649 enco |
| 15 | 921   | 93.9 | 177 | 23 AAB37020 | Human peptide enco |
| 16 | 700   | 71.4 | 190 | 21 AAB36085 | Human DSP-8. Homo  |
| 17 | 700   | 71.4 | 190 | 22 AAB73220 | Human phosphatase  |
| 18 | 697   | 71.0 | 190 | 22 AAM78868 | Human protein SEQ  |
| 19 | 642.5 | 65.5 | 199 | 22 AAM79852 | Human protein SEQ  |
| 20 | 511   | 52.1 | 198 | 20 AAY39371 | Protein phosphatas |
| 21 | 511   | 52.1 | 198 | 20 AAY28623 | Human secreted pro |
| 22 | 511   | 52.1 | 198 | 21 AAB19602 | Human dual-specifi |
| 23 | 511   | 52.1 | 198 | 21 AAY73394 | Human secreted pro |
| 24 | 504   | 51.4 | 198 | 21 AAB19620 | Dual-specificity p |
| 25 | 500   | 51.0 | 198 | 21 AAB19621 | Dual-specificity p |
| 26 | 431   | 43.9 | 151 | 22 AAM70913 | Human bone marrow  |
| 27 | 431   | 43.9 | 151 | 23 ABB40709 | Human peptide enco |
| 28 | 304   | 31.0 | 212 | 22 ABB70733 | Drosophila melanog |
| 29 | 252   | 25.7 | 367 | 23 AAU84270 | Human endometrial  |
| 30 | 252   | 25.7 | 367 | 23 ABB57320 | Mouse ischaemic co |
| 31 | 246   | 25.1 | 169 | 22 AAE06782 | Human dual-specifi |
| 32 | 246   | 25.1 | 169 | 22 AAB66438 | Human MAP-kinase p |
| 33 | 242   | 24.7 | 367 | 15 AAB63602 | MAP-kinase-phospha |
| 34 | 240   | 24.5 | 394 | 15 AAB63601 | MAP-kinase-phospha |
| 35 | 240   | 24.5 | 394 | 16 AAR78635 | Partial MAP kinase |
| 36 | 240   | 24.5 | 394 | 22 AAB76875 | Human lung tumour  |
| 37 | 240   | 24.5 | 394 | 23 AAU85530 | SV8 lung tumour p  |
| 38 | 240   | 24.5 | 476 | 22 ABB63527 | Drosophila melanog |
| 39 | 238   | 24.3 | 302 | 22 AAB76876 | Human lung tumour  |
| 40 | 238   | 24.3 | 302 | 23 AAU85531 | L551S lung tumour  |
| 41 | 238   | 24.3 | 394 | 22 AAB76878 | Human lung tumour  |
| 42 | 238   | 24.3 | 394 | 22 AAG67448 | Amino acid sequenc |
| 43 | 238   | 24.3 | 394 | 22 AAG67627 | Amino acid sequenc |
| 44 | 238   | 24.3 | 394 | 23 AAU85533 | L551S lung tumour  |
| 45 | 238   | 24.3 | 401 | 23 AAU85617 | Recombinant protei |

ALIGNMENTS

RESULT 1

AAB19008

ID AAB19008 standard; Protein; 188 AA.

XX AC AAB19008;

XX DT 08-FEB-2001 (first entry)

XX DE A human dual-specificity phosphatase 2 (DSP-2).

XX KW Human; mitogen-activated protein kinase; MAP kinase; cell survival;  
KW dual-specificity phosphatase; DSP-2; cell proliferation; DSP-2;  
KW cell differentiation; cancer; graft-versus-host disease; allergy;  
KW autoimmune disease; metabolic disease; cell cycle abnormality.

XX OS Homo sapiens.

XX PN WO200056899-A1.

XX PD 28-SEP-2000.

XX PF 22-MAR-2000; 2000WO-US07589.

XX PR 24-MAR-1999; 99US-0125957.

XX PR 16-MAR-2000; 2000US-0527376.

XX PA (CEPT-) CEPTYR INC.

XX PI Luche RM, Wei B;

XX WPI; 2000-638264/61.

XX DR N-PSDB; AAA96619.

XX XX

PT Novel dual-specificity mitogen activated protein (MAP) kinase  
PT phosphatase polypeptide useful in assays to identify agents that  
PT modulate the protein activity, which are used to treat cancer and  
PT autoimmune diseases -  
XX  
XX Claim 1; Fig 2; 51pp; English.

XX The present sequence represents a human dual-specificity mitogen-  
XX activated protein (MAP) kinase phosphatase (DSP-2). The protein has  
XX the ability to dephosphorylate an activated MAP-kinase. The DSP-2  
XX protein is used for treatment of conditions associated with cell  
XX proliferation, cell differentiation, and cell survival. The DSP-2  
XX protein is useful for screening for modulators. Agents that modulate  
XX DSP-2 activity are useful for treating cancer, graft-versus-host  
XX disease, autoimmune diseases, allergies, metabolic diseases, abnormal  
XX cell growth, abnormal cell proliferation and cell cycle abnormalities.

XX SQ Sequence 188 AA;  
Query Match 100.0%; Score 981; DB 21; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5e-110;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60  
DB 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60  
QY 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120  
DB 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120  
QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQELIHYEFQFGKNTVHMVSSPVGMIPDIYK 180  
DB 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQELIHYEFQFGKNTVHMVSSPVGMIPDIYK 180  
QY 181 EVRLMIPL 188  
DB 181 EVRLMIPL 188

RESULT 2  
AAM39323  
ID AAM39323 standard; Protein; 188 AA.  
XX AC AAM39323;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 2468.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
PR 23-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB; AAI58479.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
XX  
XX Example 4; SEQ ID NO 2468; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
XX the encoded polypeptides (AAM38642-AA42213) with nootropic,  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification.

XX SQ Sequence 188 AA;  
Query Match 100.0%; Score 981; DB 22; Length 188;  
Best Local Similarity 100.0%; Pred. No. 5e-110;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60  
DB 1 MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITPMVINVSVEVNTL 60  
QY 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120  
DB 61 YEDIQYMQVPVADSPNSRLCDFDPIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM 120  
QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQELIHYEFQFGKNTVHMVSSPVGMIPDIYK 180  
DB 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWQELIHYEFQFGKNTVHMVSSPVGMIPDIYK 180  
QY 181 EVRLMIPL 188  
DB 181 EVRLMIPL 188

RESULT 3  
AAB85360  
ID AAB85360 standard; Protein; 188 AA.  
XX AC AAB85360;  
XX  
DT 17-SEP-2001 (first entry)  
XX  
XX Human phosphatase (PP) (clone ID 6205333CD1).  
XX  
KW Phosphatase; PP; Alzheimer's disease; Huntington's disease; human;  
KW immune disorder; cancer; nootropic; immunomodulatory; cytostatic;  
KW neuroprotective; anticonvulsant.  
OS Homo sapiens.  
XX  
XX WO200153469-A2.  
XX

PD 26-JUL-2001.  
 XX 18-JAN-2001; 2001WO-US02088.  
 XX 21-JAN-2000; 2000US-0177719.  
 PR 28-JAN-2000; 2000US-0178988.  
 PR 25-FEB-2000; 2000US-0184959.  
 PR 17-MAR-2000; 2000US-0190142.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Bandman O, Tang YT, Azimzai Y, Yue H, Baughn MR, Hillman JL;  
 PI Lal P, Wang E, Gandhi AR, Policky JL, Mathur P;  
 XX WPI; 2001-451905/48.  
 DR N-PSDB; AAH22963.  
 XX Novel phosphatases useful for treating diseases associated with  
 PT decreased expression of functional phosphatases, e.g., Alzheimer's  
 PT disease, Huntington's disease, immune disorders, and cancers  
 XX Claim 1; Page 95-96; 103pp; English.  
 XX The invention provides human phosphatases (pp) and polynucleotides  
 CC encoding the phosphatases. The polypeptides can be expressed by standard  
 CC recombinant methodology. The pp are useful for treating a disease or  
 CC condition associated with decreased expression of functional  
 CC phosphatases. Compositions containing agonists or antagonists of pp may  
 CC be used to treat a disease associated with decreased expression or  
 CC overexpression of pp, respectively. Such diseases may include  
 CC Alzheimer's disease, Huntington's disease, immune disorders, and  
 CC cancers. The present sequence represents a human phosphatase.  
 XX Sequence 188 AA;  
 SQ  
 Query Match 100.0%; Score 981; DB 22; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 5e-110;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WTAPSCAFVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60  
 DB 1 WTAPSCAFVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60  
 QY 61 YEDIQMVQPVADSPNSRLCDFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLM 120  
 DB 61 YEDIQMVQPVADSPNSRLCDFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLM 120  
 QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180  
 DB 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180  
 QY 181 EVRLMIPL 188  
 DB 181 EVRLMIPL 188  
 RESULT 4  
 AAB73221  
 ID AAB73221 standard; Protein; 188 AA.  
 XX AAB73221;  
 XX 11-MAY-2001 (first entry)  
 DT Human phosphatase AA915932\_h.  
 XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;  
 KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;  
 KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;  
 KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;  
 KW schizophrenia; hamartoma.  
 XX Homo sapiens.  
 OS

XX WO200112819-A2.  
 PN 22-FEB-2001.  
 XX 11-AUG-2000; 2000WO-US22158.  
 PF 13-AUG-1999; 99US-0149005.  
 XX (SUGE-) SUGEN INC.  
 PA Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;  
 PI WPI; 2001-211226/21.  
 XX N-PSDB; AAF63573.  
 DR New protein phosphatase polypeptide for diagnosing and treating  
 PT phosphatase related disorders such as cancer, schizophrenia, cardiac  
 PT dysfunction and/or vascular disorders  
 XX Claim 6; Fig 5; 138pp; English.  
 PS The present invention relates to phosphatase proteins and coding  
 XX sequences. The present sequence is one such phosphatase. Phosphatases are  
 CC enzymes that catalyze the dephosphorylation of proteins modified by  
 CC phosphorylation of serine, threonine or tyrosine residues. The  
 CC phosphatases are useful for treating a variety of diseases: for example  
 CC cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,  
 CC synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,  
 CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,  
 CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological  
 CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,  
 CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,  
 CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan  
 CC Zonana syndrome, schizophrenia and hamartomas.  
 XX Sequence 188 AA;  
 SQ  
 Query Match 100.0%; Score 981; DB 22; Length 188;  
 Best Local Similarity 100.0%; Pred. No. 5e-110;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WTAPSCAFVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60  
 DB 1 WTAPSCAFVQRPQPSVGLSGLSITKSLYISNGVAANNKMLSSNQITWVINSVEVNTL 60  
 QY 61 YEDIQMVQPVADSPNSRLCDFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLM 120  
 DB 61 YEDIQMVQPVADSPNSRLCDFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLM 120  
 QY 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180  
 DB 121 KYHMSLLDAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180  
 QY 181 EVRLMIPL 188  
 DB 181 EVRLMIPL 188  
 RESULT 5  
 AAU22098  
 ID AAU22098 standard; Protein; 192 AA.  
 XX AAU22098;  
 XX 17-DEC-2001 (first entry)  
 DT Human cardiovascular system antigen polypeptide SEQ ID No 872.  
 DE Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
 XX chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
 KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
 KW cerebroprotective; nootropic; antitubacterial; virucide; fungicide; cancer;  
 KW

|    |                                                                         |    |              |                |
|----|-------------------------------------------------------------------------|----|--------------|----------------|
| KW | ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  | PR | 14-SEP-2000; | 2000US-0233063 |
| KW | hyperproliferative disorder; breast; liver; cardiovascular disorder;    | PR | 14-SEP-2000; | 2000US-0233064 |
| KW | cerebrovascular disorder; nervous system disorder; bacterial infection; | PR | 14-SEP-2000; | 2000US-0233065 |
| KW | fungal infection; viral infection; ocular disorder; endocrine disorder; | PR | 21-SEP-2000; | 2000US-0234223 |
| KW | gastrointestinal disorder; renal disorder; respiratory disorder;        | PR | 21-SEP-2000; | 2000US-0234274 |
| KW | wound healing; skin aging; organ transplantation; tissue regeneration;  | PR | 25-SEP-2000; | 2000US-0234997 |
| KW | anti-infertility.                                                       | PR | 25-SEP-2000; | 2000US-0234998 |
| OS | Homo sapiens.                                                           | PR | 26-SEP-2000; | 2000US-0235484 |
| XX |                                                                         | PR | 27-SEP-2000; | 2000US-0235834 |
| XX |                                                                         | PR | 27-SEP-2000; | 2000US-0235836 |
| PN |                                                                         | PR | 29-SEP-2000; | 2000US-0236327 |
| XX |                                                                         | PR | 29-SEP-2000; | 2000US-0236367 |
| XX |                                                                         | PR | 29-SEP-2000; | 2000US-0236368 |
| PF |                                                                         | PR | 29-SEP-2000; | 2000US-0236369 |
| PF |                                                                         | PR | 29-SEP-2000; | 2000US-0236370 |
| XX |                                                                         | PR | 02-OCT-2000; | 2000US-0236802 |
| XX |                                                                         | PR | 02-OCT-2000; | 2000US-0237037 |
| XX |                                                                         | PR | 02-OCT-2000; | 2000US-0237038 |
| PR |                                                                         | PR | 02-OCT-2000; | 2000US-0237039 |
| PR |                                                                         | PR | 02-OCT-2000; | 2000US-0237040 |
| PR |                                                                         | PR | 13-OCT-2000; | 2000US-0239935 |
| PR |                                                                         | PR | 13-OCT-2000; | 2000US-0239937 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0240960 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0241221 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0241785 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0241786 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0241787 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0241808 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0241809 |
| PR |                                                                         | PR | 20-OCT-2000; | 2000US-0241826 |
| PR |                                                                         | PR | 01-NOV-2000; | 2000US-0244617 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246474 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246475 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246476 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246477 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246478 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246523 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246524 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246525 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246526 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246527 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246528 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246532 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246609 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246610 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246611 |
| PR |                                                                         | PR | 08-NOV-2000; | 2000US-0246613 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249207 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249208 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249209 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249210 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249211 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249212 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249213 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249214 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249215 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249216 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249217 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249218 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249244 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249245 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249264 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249265 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249297 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249299 |
| PR |                                                                         | PR | 17-NOV-2000; | 2000US-0249300 |
| PR |                                                                         | PR | 01-DEC-2000; | 2000US-0250160 |
| PR |                                                                         | PR | 01-DEC-2000; | 2000US-0250391 |
| PR |                                                                         | PR | 05-DEC-2000; | 2000US-0251030 |
| PR |                                                                         | PR | 05-DEC-2000; | 2000US-0251988 |
| PR |                                                                         | PR | 05-DEC-2000; | 2000US-0256719 |
| PR |                                                                         | PR | 06-DEC-2000; | 2000US-0251479 |
| PR |                                                                         | PR | 08-DEC-2000; | 2000US-0251856 |



PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251899.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-451930/48.  
 DR N-PSDB; AAS35372.  
 XX  
 PT New cardiovascular system related polynucleotides and polypeptides,  
 PT useful for diagnosing, treating and/or preventing disorders of the  
 PT cardiovascular system -  
 XX  
 PS Claim 11; SEQ ID NO 872; 674pp; English.  
 XX  
 CC Sequences AAU21852-AAU22466 represent the cardiovascular system antigen  
 CC polypeptides of the invention. Cardiovascular system antigens and their  
 CC associated polynucleotides are useful in the diagnosis, treatment and  
 CC prevention of various types of disorders in e.g. humans, mice, rabbits,  
 CC goats, horses, cats, dogs, chickens or sheep. A pathological condition  
 CC can be determined by detecting the presence or absence of a mutation in a  
 CC cardiovascular system antigen polynucleotide. The treatable disorders  
 CC include autoimmune diseases such as rheumatoid arthritis,  
 CC hyperproliferative disorders such as neoplasms of the breast or liver,  
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular  
 CC disorders such as cerebral ischaemia, nervous system disorders and fungi.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi.  
 CC ocular disorders such as corneal infection, endocrine disorders such as  
 CC premature labour and infertility, gastrointestinal disorders such as  
 CC Crohn's disease, renal disorders such as glomerulonephritis and  
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can  
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,  
 CC to maintain organs before transplantation, to regenerate tissues and in  
 CC chemotaxis.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

Query Match 100.0%; Score 981; DB 22; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-110;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTAPSCAFPVQFRQPSVSGLSQITKSLYISNGVAANNKMLSSNQITWVINVSEVNTL 60  
 Db 5 MTAPSCAFPVQFRQPSVSGLSQITKSLYISNGVAANNKMLSSNQITWVINVSEVNTL 64  
 Qy 61 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVEMKQRTLLHCAAGYSRAALCLAYLM 120  
 Db 65 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVEMKQRTLLHCAAGYSRAALCLAYLM 124  
 Qy 121 KYHMSLLDAHTWTKSCRPPIRPNISGFWQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180  
 Db 125 KYHMSLLDAHTWTKSCRPPIRPNISGFWQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 184  
 Qy 181 EVRLMIPL 188  
 Db 185 EVRLMIPL 192

RESULT 6  
 AA41109  
 ID AA41109 standard; Protein; 192 AA.  
 XX  
 AC AA41109;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 6040.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX Homo sapiens.  
 OS  
 XX WO200153312-A1.  
 PN  
 XX 26-JUL-2001.  
 PD  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 PF  
 XX 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR N-PSDB; AAI60265.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Example 2; SEQ ID NO 6040; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemia and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

Sequence 192 AA;

Query Match 100.0%; Score 981; DB 22; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-110;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTAPSCAFPVQFRQPSVSGLSQITKSLYISNGVAANNKMLSSNQITWVINVSEVNTL 60  
 Db 5 MTAPSCAFPVQFRQPSVSGLSQITKSLYISNGVAANNKMLSSNQITWVINVSEVNTL 64  
 Qy 61 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVEMKQRTLLHCAAGYSRAALCLAYLM 120  
 Db 65 YEDIOYMQVPVADSPNSRLCDFDFDIADHHSVEMKQRTLLHCAAGYSRAALCLAYLM 124  
 Qy 121 KYHMSLLDAHTWTKSCRPPIRPNISGFWQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 180  
 Db 125 KYHMSLLDAHTWTKSCRPPIRPNISGFWQLIHYEFQFGKNTVHMVSPVGMIPDIYEK 184

QY 181 EVRLMPL 188  
 Db 185 EVRLMPL 192  
 RESULT 7  
 AAU21820  
 ID AAU21820 standard; Protein; 192 AA.  
 AC AAU21820;  
 XX  
 AC  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Novel human neoplastic disease associated polypeptide #253.  
 XX  
 KW Human; neoplastic disease associated polypeptide; cancer;  
 KW hyperproliferative disorder; neural disorder; immune system disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW neuroprotective; cytostatic; anti inflammatory; vasotropic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20015163-A1.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01358.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 06-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246533.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
DR WPI; 2001-465558/50.  
DR N-PSDB; AAS35019.  
XX  
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to  
PT diagnose diseases or disorders associated with aberrant expression or  
PT activity of polypeptides, and for treating cancers, rheumatoid  
PT arthritis  
XX  
XX  
PS Claim 11; SEQ ID No 547; 687pp; English.  
XX  
CC The present invention relates to the isolation of novel human neoplastic  
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA  
CC sequences encoding for these polypeptides. The sequences of the  
CC invention are useful in the diagnosis, treatment, prevention and/or  
CC prognosis of disorders involving neoplastic disease such as  
CC hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder  
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar  
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may  
CC also be useful for treating other disorders such as neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC and renal disorders. The polynucleotide sequences of the invention are  
CC also useful in gene therapy. AA021568-AA021851 represent the novel human  
CC neoplastic disease associated polypeptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 192 AA;  
Query Match 100.0%; Score 981; DB 22; Length 192;  
Best Local Similarity 100.0%; Pred. No. 5.2e-110;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MTPASCAFPVQFROPSPVSGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVVNTL 60  
D6 5 MTPASCAFPVQFROPSPVSGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVVNTL 64  
QY 61 YEDIOYMQVPVADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLM 120  
D6 65 YEDIOYMQVPVADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLM 124  
QY 121 KYHAMSLLDARTWTKSCRPIIRPNSGFWEQLIHYEFOLFGKNTVHMVSSPVGMLPDIYK 180  
D6 125 KYHAMSLLDARTWTKSCRPIIRPNSGFWEQLIHYEFOLFGKNTVHMVSSPVGMLPDIYK 184  
QY 181 EVRLMIPL 188  
D6 181 EVRLMIPL 188

DB 185 EVRLMIPL 192  
RESULT 8  
ID ABB29053 standard; Peptide; 177 AA.  
XX AC ABB29053;  
XX DT 01-FEB-2002 (first entry)  
XX Peptide #1704 encoded by breast cell single exon nucleic acid probe.  
DE Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer.  
XX Homo sapiens.  
XX WO200157271-A2.  
XX 09-AUG-2001.  
XX 30-JAN-2001; 2001WO-US00662.  
XX 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-496933/54.  
XX New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes -  
XX Claim 27; SEQ ID NO 12021; 327pp + sequence listing; English.  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a peptide encoded by a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 177 AA;  
Query Match 93.9%; Score 921; DB 22; Length 177;  
Best Local Similarity 100.0%; Pred. No. 8.2e-103;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 FRQPSVSGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVVNTLIEDIOYMQVPV 71  
D6 1 FRQPSVSGLSQITKSLYISNGVAANKMLSSNQITWVINSVEVVNTLIEDIOYMQVPV 60

QY 72 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 131  
Db 61 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 120  
QY 132 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 188  
Db 121 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 177

RESULT 9  
ABB34215  
ID ABB34215 standard; Peptide; 177 AA.  
AC ABB34215;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #1721 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver.  
XX  
PS Claim 27; SEQ ID NO 26850; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 177 AA;

Query Match 93.9%; Score 921; DB 22; Length 177;  
Best Local Similarity 100.0%; Pred. No. 8.2e-103;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVEVNTLYEDIQYMQVPV 71  
Db 1 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVEVNTLYEDIQYMQVPV 60  
QY 72 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 131  
Db 61 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 120  
QY 132 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 188  
Db 121 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 177

QY 132 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 188  
Db 121 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 177

RESULT 10  
AAM55008  
ID AAM55008 standard; Protein; 177 AA.  
XX  
AC AAM55008;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27113.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483446/52.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains.  
XX  
PS Example 4; SEQ ID NO: 27113; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
SQ Sequence 177 AA;

Query Match 93.9%; Score 921; DB 22; Length 177;  
Best Local Similarity 100.0%; Pred. No. 8.2e-103;  
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVEVNTLYEDIQYMQVPV 71  
Db 1 FROPVSGLSQITKSLYISNGVAANNKMLSSNQITMTWINSVVEVNTLYEDIQYMQVPV 60  
QY 72 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 131  
Db 61 ADSPNSRLCDFDPTADHHSVEMKQGRLLHCAAGVSRSAALCLAYLMKYHMSLLDAH 120  
QY 132 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 188  
Db 121 TWTSCRPPIIRPNSGFWQLIHYEFQFGKNTVHMVSSPVGMPDIIYEKEVRLMPL 177

```
RESULT 11
AAM67394
ID AAM67394 standard; Protein; 177 AA.
XX
AC AAM67394;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27700.
XX
DE Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 27700; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 177 AA;
Query Match 93.9%; Score 921; DB 22; Length 177;
Best Local Similarity 100.0%; Pred. No. 8.2e-103;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVVNTLYEDIQYMQVPV 71
DB 1 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVVNTLYEDIQYMQVPV 60
QY 72 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 131
DB 61 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 120
QY 132 TWTKSCRPIIRPNNSGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 188
DB 121 TWTKSCRPIIRPNNSGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 177
RESULT 12
AAM15224
ID AAM15224 standard; Protein; 177 AA.
XX
AC AAM15224;
XX
DT 12-OCT-2001 (first entry)
```

```
XX
DE Peptide #1658 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID No 20050; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI0068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 177 AA;
Query Match 93.9%; Score 921; DB 22; Length 177;
Best Local Similarity 100.0%; Pred. No. 8.2e-103;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVVNTLYEDIQYMQVPV 71
DB 1 FROPVSGLSQITKSLYISNGVAANKMLSSNQITMWINSVVNTLYEDIQYMQVPV 60
QY 72 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 131
DB 61 ADSPNSRLCDFDPIADHHSVEMKOGRTLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 120
QY 132 TWTKSCRPIIRPNNSGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 188
DB 121 TWTKSCRPIIRPNNSGFWEQLIHYEFQFGKNTVHMVSSPVGMPIDIYEKEVRLMPL 177
RESULT 13
AAM27685
ID AAM27685 standard; Protein; 177 AA.
XX
AC AAM27685;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #1722 encoded by probe for measuring placental gene expression.
```

XX Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder.

OS Homo sapiens.

XX WO200157272-A2.

XX PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

XX Claim 27; SEQ ID No 27954; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;

CC see AA13135-AA157546). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for producing a microarray for

CC predicting, measuring and displaying gene expression in samples derived

CC from human placenta. The probes are useful for antenatal diagnosis of

CC human genetic disorders.

XX Sequence 177 AA;

Qy 12 FRQPSVGLSQITKSLYISNGVAANKLMLSSNQITWVINSVEVNTLYEDIQMVQPV 71

Db 1 FRQPSVGLSQITKSLYISNGVAANKLMLSSNQITWVINSVEVNTLYEDIQMVQPV 60

Qy 72 ADSPNSRLCDFDPIADHIHSHVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 131

Db 61 ADSPNSRLCDFDPIADHIHSHVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 120

Qy 132 TWTKSCRIIPRNSGFWEQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 188

Db 121 TWTKSCRIIPRNSGFWEQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 177

RESULT 14

AA02967

ID AA02967 standard; Protein; 177 AA.

XX AA02967;

XX 09-OCT-2001 (first entry)

XX Peptide #1649 encoded by probe for measuring breast gene expression.

XX Probe; human; breast disease; breast cancer; development disorder;

XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

OS Homo sapiens.

XX WO200157270-A2.

PN

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US00661.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression

PT in a human breast -

XX Claim 27; SEQ ID No 11707; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes

CC (see AA100010-AA110067). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for measuring human gene expression in

CC a human breast sample, where the probe hybridizes at high stringency to a

CC nucleic acid expressed in the human breast. The probes are useful for

CC predicting, diagnosing, grading, staging, monitoring and prognosing

CC diseases of the human breast, particularly those diseases with polygenic

CC aetiology. The diseases include: breast cancer, disorders of development,

CC inflammatory diseases of the breast, fibrocystic changes, proliferative

CC breast disease and non-carcinoma tumours.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 177 AA;

Qy 12 FRQPSVGLSQITKSLYISNGVAANKLMLSSNQITWVINSVEVNTLYEDIQMVQPV 71

Db 1 FRQPSVGLSQITKSLYISNGVAANKLMLSSNQITWVINSVEVNTLYEDIQMVQPV 60

Qy 72 ADSPNSRLCDFDPIADHIHSHVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 131

Db 61 ADSPNSRLCDFDPIADHIHSHVENKQGRITLLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 120

Qy 132 TWTKSCRIIPRNSGFWEQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 188

Db 121 TWTKSCRIIPRNSGFWEQLIHYEFQFGKNTVHMVSSPVGMIPDIYEKEVRLMIPL 177

RESULT 15

ABG37020

ID ABG37020 standard; Peptide; 177 AA.

XX ABG37020;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 26685.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

XX familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
primary ciliary dyskinesia; pulmonary hypertension;  
hyaline membrane disease.

**Homo sapiens.**

WO200186003-A2.

15-NOV-2001.

30-JAN-2001; 2001WO-US00665.

04-FEB-2000; 2000US-180312P.

20 MAY 2000; 2000US-207430E;  
30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366;  
21-SEP-2000; 2000US-234687P;

27-SEP-2000; 2000US-236339F;  
04-OCT-2000; 2000GB-0024263;

(MOL.F.-) MOLECULAR DYNAMICS T

Penn SG, Hanzel DK, Chen W

WPT: 2002-114183/15

Spatially-v-addressable set of si

measure gene expression in human lung s

Claim 27; SEQ ID No

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

Claim 27; SEQ ID No 26685; 634pp; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridize at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

Sequence 177 AA;

|    | Query Match           | 93.9%;                                                         | Score 921;          | DB 23;    | Length 177; |
|----|-----------------------|----------------------------------------------------------------|---------------------|-----------|-------------|
|    | Best Local Similarity | 100.0%;                                                        | Prod. No. 8.2e-103; |           |             |
|    | Matches 177;          | Conservative 0;                                                | Mismatches 0;       | Indels 0; | Gaps 0;     |
| Qy | 12                    | PROPSVSGLSQITKSLYISNGVAAANKMLMLSSNQITMWINVSVVWNTLYEDIQYMQVPV   | 71                  |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
| Db | 1                     | FRQPSVSGLSQITKSLYISNGVAAANKMLMLSSNQITMWINVSVVWNTLYEDIQYMQVPV   | 60                  |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
| Qy | 72                    | ADSPNSRLCDDFFDIADHIIHSVENMKQGRTHLHCAAGYSRSAALCLAYLMKYHAMSLLDAH | 131                 |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
| Db | 61                    | ADSPNSRLCDDFFDIADHIIHSVENMKQGRTHLHCAAGYSRSAALCLAYLMKYHAMSLLDAH | 120                 |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
| Qy | 132                   | TWTKSCRPRIIPNSGFWQLIHYEFQFGKNTVHMVSSPGMIPDIIYEKEVRLMIPL        | 188                 |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
| Db | 121                   | TWTKSCRPRIIPNSGFWQLIHYEFQFGKNTVHMVSSPGMIPDIIYEKEVRLMIPL        | 177                 |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |
|    |                       |                                                                |                     |           |             |

Search completed: April 21, 2003, 18:32:34  
Job time : 44 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 18:35:04 ; Search time 114 seconds  
(without alignments)  
124.637 Million cell updates/sec

Title: US-09-527-376-2  
Perfect score: 981  
Sequence: 1 MTAPSCAFFVQFRQPSVGL.....SPVGMIPDIYEKVRMLMPL 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 981   | 100.0       | 192    | 9     | US-10-091-504-872   |
| 2          | 981   | 100.0       | 192    | 10    | US-09-764-869-872   |
| 3          | 921   | 93.9        | 177    | 10    | US-09-864-761-42750 |
| 4          | 700   | 71.4        | 190    | 12    | US-10-044-205A-44   |
| 5          | 697   | 71.0        | 190    | 12    | US-10-044-205A-42   |
| 6          | 431   | 43.9        | 151    | 10    | US-09-864-761-46915 |
| 7          | 252   | 25.7        | 367    | 10    | US-09-919-497-60    |
| 8          | 246   | 25.1        | 154    | 9     | US-09-955-732-8     |
| 9          | 246   | 25.1        | 154    | 10    | US-09-964-277-9     |
| 10         | 246   | 25.1        | 169    | 10    | US-09-775-925-28    |
| 11         | 240   | 24.5        | 169    | 10    | US-09-847-519A-13   |
| 12         | 240   | 24.5        | 394    | 9     | US-09-736-457-805   |
| 13         | 240   | 24.5        | 394    | 9     | US-09-902-941-805   |
| 14         | 240   | 24.5        | 394    | 9     | US-09-849-626-805   |
| 15         | 240   | 24.5        | 394    | 9     | US-10-017-754-805   |
| 16         | 238   | 24.3        | 302    | 9     | US-09-736-457-806   |
| 17         | 238   | 24.3        | 302    | 9     | US-09-902-941-806   |
| 18         | 238   | 24.3        | 302    | 9     | US-09-849-626-806   |
| 19         | 238   | 24.3        | 302    | 9     | US-10-017-754-806   |

|    |       |      |     |    |                    |                    |
|----|-------|------|-----|----|--------------------|--------------------|
| 20 | 238   | 24.3 | 394 | 9  | US-09-736-457-827  | Sequence 827, App  |
| 21 | 238   | 24.3 | 394 | 9  | US-09-902-941-827  | Sequence 827, App  |
| 22 | 238   | 24.3 | 394 | 9  | US-09-849-626-827  | Sequence 827, App  |
| 23 | 238   | 24.3 | 394 | 9  | US-10-017-754-827  | Sequence 827, App  |
| 24 | 238   | 24.3 | 401 | 9  | US-09-902-941-1917 | Sequence 1917, App |
| 25 | 238   | 24.3 | 401 | 9  | US-09-849-626-1917 | Sequence 1917, App |
| 26 | 238   | 24.3 | 401 | 9  | US-10-017-754-1917 | Sequence 1917, App |
| 27 | 230   | 23.4 | 235 | 10 | US-09-963-204-2    | Sequence 2, Appli  |
| 28 | 230   | 23.4 | 289 | 10 | US-09-963-204-12   | Sequence 12, Appl  |
| 29 | 230   | 23.4 | 329 | 10 | US-09-963-204-22   | Sequence 22, Appl  |
| 30 | 227   | 23.1 | 154 | 9  | US-09-955-732-9    | Sequence 9, Appli  |
| 31 | 227   | 23.1 | 154 | 10 | US-09-964-277-10   | Sequence 10, Appl  |
| 32 | 227   | 23.1 | 169 | 10 | US-09-775-925-29   | Sequence 29, Appl  |
| 33 | 227   | 23.1 | 169 | 10 | US-09-847-519A-14  | Sequence 14, Appl  |
| 34 | 221   | 22.5 | 154 | 9  | US-09-955-732-10   | Sequence 10, Appl  |
| 35 | 221   | 22.5 | 154 | 10 | US-09-964-277-11   | Sequence 11, Appl  |
| 36 | 221   | 22.5 | 171 | 10 | US-09-775-925-30   | Sequence 30, Appl  |
| 37 | 221   | 22.5 | 171 | 10 | US-09-847-519A-15  | Sequence 15, Appl  |
| 38 | 220   | 22.4 | 173 | 10 | US-09-816-494-7    | Sequence 7, Appli  |
| 39 | 220   | 22.4 | 173 | 10 | US-09-815-419-4    | Sequence 4, Appli  |
| 40 | 219.5 | 22.4 | 170 | 10 | US-09-775-925-26   | Sequence 26, Appl  |
| 41 | 219.5 | 22.4 | 170 | 10 | US-09-847-519A-11  | Sequence 11, Appl  |
| 42 | 217   | 22.1 | 217 | 10 | US-09-815-419-2    | Sequence 2, Appli  |
| 43 | 213.5 | 21.8 | 172 | 10 | US-09-816-494-8    | Sequence 8, Appli  |
| 44 | 213.5 | 21.8 | 172 | 10 | US-09-815-419-5    | Sequence 5, Appli  |
| 45 | 211   | 21.5 | 155 | 9  | US-09-955-732-6    | Sequence 6, Appli  |

ALIGNMENTS

RESULT 1:  
US-10-091-504-872  
; Sequence 872, Application US/10091504  
; Publication NO. US20030059908A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007C1  
; CURRENT APPLICATION NUMBER: US/10/091,504  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 2442  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 872  
; LENGTH: 192  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-091-504-872

Query Match 100.0%; Score 981; DB 9; Length 192;  
Best Local Similarity 100.0%; Pred. No. 9.3e-101;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|                   |     |                                                           |     |
|-------------------|-----|-----------------------------------------------------------|-----|
| QY                | 1   | MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITWINSVEVVNTL   | 60  |
| DB                | 5   | MTAPSCAFFVQFRQPSVGLSQITKSLYISNGVAANKMLSSNQITWINSVEVVNTL   | 64  |
| QY                | 61  | YEDIQMVQPVADSPNSRLCDFDPDIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM | 120 |
| DB                | 65  | YEDIQMVQPVADSPNSRLCDFDPDIADHHSVEMKQGRLLHCAAGVSRSAALCLAYLM | 124 |
| QY                | 121 | KYHMSLLDAHTWTKSCRIIPRNSGFWQLIHYEQFGKNTVHMVSSPVGMPDIYEK    | 180 |
| DB                | 125 | KYHMSLLDAHTWTKSCRIIPRNSGFWQLIHYEQFGKNTVHMVSSPVGMPDIYEK    | 184 |
| QY                | 181 | EVRLMIPL 188                                              |     |
| DB                | 185 | EVRLMIPL 192                                              |     |
| RESULT 2          |     |                                                           |     |
| US-09-764-869-872 |     |                                                           |     |







[illegible]

```

RESULT 11
US-09-847-519A-13
; Sequence 13, Application US/09847519A
; Patent NO. US20020102693A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-14 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125,422
; CURRENT APPLICATION NUMBER: US/09/847,519A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 13
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-847-519A-13

```

|                              |                |                                                              |         |             |
|------------------------------|----------------|--------------------------------------------------------------|---------|-------------|
| Query Match                  | 25.1%          | Score 246;                                                   | DB 10;  | Length 169; |
| Best Local Similarity        | 36.4%;         | Pred. No. 1.7e-19;                                           |         |             |
| Matches 56; Conservative 27; | Mismatches 67; | Indels 4;                                                    | Gaps 1; |             |
| QY                           | 5              | SCAPVQFRQPVSGLSQITKSLIISNCVAANKMLSSNQITWVINVSVEVVNTLYEDI     | 64      |             |
| Dd                           | 16             | SCSTPLY-----DQGSPVELLPFLYLGSAHASKRMDLDAIGTALINYSANCPHFEGHY   | 71      |             |
| QY                           | 65             | QYMVPVADSPNSRLCDDFFDIADHIHSVMKQGTLLHCAGVSRSASALCLAYLMKYHA    | 124     |             |
| Dd                           | 72             | QYKSIPVEDNHKAIDISWFNEAIDFIDSIKNAGGVFVHCQAGISRSTATICLAYLMRTNR | 131     |             |
| QY                           | 125            | MSLDATHTWKSCRPIIRNPNSGFWEOLIHFEQL                            | 158     |             |
| Dd                           | 132            | VKLDEAEFVKORRSIIISPNSFMGLOLFESOV                             | 165     |             |

RESULT 12  
US-09-736-457-805  
Sequence 805, Application US/09736457  
Patent No. US20020168637A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
APPLICANT: Wang, AiJun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITILE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C15  
CURRENT APPLICATION NUMBER: US/09/736,457  
CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 1864  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 805  
LENGTH: 394  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-736-457-805

|    | Query Match           | 24.5%                                                          | score 240;                                   | DB 9;     | Length 394; |
|----|-----------------------|----------------------------------------------------------------|----------------------------------------------|-----------|-------------|
|    | Best Local Similarity | 33.9%;                                                         | pred. No. 2.5e-18;                           |           |             |
|    | Matches 57;           | Conservative 32;                                               | Mismatches 75;                               | Indels 4; | Gaps 1;     |
| Qy | 5                     | SCAPVQRFPQSGLS                                                 | OITKSLYISNGVAANKMLSSNOITWTVINVSVVVNTLYEDI    | 64        |             |
| Dd | 185                   | SGTFLPHDQE---                                                  | GPVEILPFLYGSAHYAARDMDLDALGITALLNVSDCPNHFEGHY | 240       |             |
| Qy | 65                    | GYMVPPVADSPNSRLCDEFFDIADHHSVENKMGGOTLLHCAAAGYSRSRAALCLAYLMKYHA | 124                                          |           |             |
| Dd | 241                   | QYKCIPVEDNHKAIDISSWFMEAEIYDAVKDCRGVLVHCQAIGTSRATICLAYLMMKKR    | 300                                          |           |             |
| Qy | 125                   | MSLDAHTTKTCRPRIIRNPISGFWEQLTHYBFOLFGRKNVTVMHVSPVG              | 172                                          |           |             |
| Dd | 301                   | VRLAEAFEFVKORSIIISPNSFMCOLLOFESOVLATSCAAEAASPSG                | 348                                          |           |             |

```

RESULT 13
US-09-902-941-805
; Sequence 805, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF LUNG CANCER
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-941-805

```

[illegible]

```

RESULT 14
US -9-849-626-805
; Sequence 805, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaityana
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, AiJun
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Tongtong

```

```
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-626-805

Query Match      24.5%; Score 240; DB 9; Length 394;
Best Local Similarity 33.9%; Pred. No. 2.5e-18;
Matches 57; Conservative 32; Mismatches 75; Indels 4; Gaps 1;

Qy 5 SCAPPVQRPQPSVGLSQITKSLYISNGVAANNKMLSSNOITMVINVSVEVNTLYEDI 64
   || || : : : : || : : || : : || : : || : : || : : || : : ||
Db 185 SCGTPLDHQE---GPVEILPFLVLSAYHAARDMLDALGITALLNVSSDCPNHFEHY 240

Qy 65 QYMQVPVADSPNSRLCDFDPIADHHSVENMKQGRFTLLHCAAGVSRSAALCLAYILMKYHA 124
   || || : : : : || : : || : : || : : || : : || : : || : : ||
Db 241 QYKICPVEDNHKADISSWFMEAEIYDAVKDCRGRVLVHCQAGISRATICLAYILMKKR 300

Qy 125 MSLDHAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSSPVG 172
   || || : : : : || : : || : : || : : || : : || : : || : : ||
Db 301 VRLEAEFEFVKQRRIISPNFSFMGQLLOFESQVLATSCAAEAASPSG 348

RESULT 15
US-10-017-754-805
; Sequence 805, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 805
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-754-805

Query Match      24.5%; Score 240; DB 9; Length 394;
Best Local Similarity 33.9%; Pred. No. 2.5e-18;
Matches 57; Conservative 32; Mismatches 75; Indels 4; Gaps 1;

Qy 5 SCAPPVQRPQPSVGLSQITKSLYISNGVAANNKMLSSNOITMVINVSVEVNTLYEDI 64
   || || : : : : || : : || : : || : : || : : || : : || : : ||
Db 185 SCGTPLDHQE---GPVEILPFLVLSAYHAARDMLDALGITALLNVSSDCPNHFEHY 240

Qy 65 QYMQVPVADSPNSRLCDFDPIADHHSVENMKQGRFTLLHCAAGVSRSAALCLAYILMKYHA 124
   || || : : : : || : : || : : || : : || : : || : : || : : ||
Db 241 QYKICPVEDNHKADISSWFMEAEIYDAVKDCRGRVLVHCQAGISRATICLAYILMKKR 300
```

```
Qy 125 MSLDHAHTWTKSCRPIIRPNSGFWEQLIHYEFQFGKNTVHMVSSPVG 172
   || || : : : : || : : || : : || : : || : : || : : || : : ||
Db 301 VRLEAEFEFVKQRRIISPNFSFMGQLLOFESQVLATSCAAEAASPSG 348

Search completed: April 21, 2003, 18:41:45
Job time : 115 secs
```

| Result No. | Query % |       |        | DB | ID     | Description        |
|------------|---------|-------|--------|----|--------|--------------------|
|            | Score   | Match | Length |    |        |                    |
| 1          | 338     | 34.5  | 272    | 2  | T18915 | hypothetical prote |
| 2          | 275     | 28.0  | 226    | 2  | T21380 | hypothetical prote |
| 3          | 252     | 25.7  | 367    | 1  | S20900 | dual specificity p |
| 4          | 252     | 25.7  | 367    | 1  | S24411 | dual specificity p |
| 5          | 252     | 25.7  | 367    | 2  | S52265 | dual specificity p |
| 6          | 238     | 24.3  | 303    | 2  | T46405 | hypothetical prote |
| 7          | 236     | 24.1  | 393    | 2  | A56947 | dual specificity p |
| 8          | 236     | 24.1  | 394    | 2  | A56115 | dual specificity p |
| 9          | 233     | 23.8  | 619    | 2  | T15969 | hypothetical prote |
| 10         | 231.5   | 23.6  | 186    | 2  | T16056 | hypothetical prote |
| 11         | 223.5   | 22.8  | 365    | 2  | T32494 | hypothetical prote |
| 12         | 221     | 22.5  | 384    | 1  | I38890 | dual specificity p |
| 13         | 208     | 21.2  | 314    | 2  | B57126 | dual specificity p |
| 14         | 197     | 20.1  | 314    | 1  | A57126 | dual specificity p |
| 15         | 180     | 18.3  | 223    | 2  | I43365 | protein tyrosine p |
| 16         | 170     | 17.3  | 142    | 2  | T03074 | dual specificity p |
| 17         | 165.5   | 16.9  | 292    | 2  | S41012 | hypothetical prote |
| 18         | 164.5   | 16.8  | 330    | 2  | T39698 | protein tyrosine p |
| 19         | 162.5   | 16.6  | 185    | 1  | A47196 | dual specificity p |
| 20         | 162     | 16.5  | 866    | 2  | F88481 | dual specificity p |
| 21         | 160     | 16.3  | 209    | 1  | S48459 | protein C16A3.1 [i |
| 22         | 158     | 16.1  | 771    | 2  | T47666 | probable dual spec |
| 23         | 157     | 16.0  | 205    | 2  | I49364 | phosphatase-like p |
| 24         | 155     | 15.8  | 489    | 1  | S58725 | protein tyrosine p |
| 25         | 151     | 15.4  | 171    | 1  | B47452 | dual specificity p |
| 26         | 147     | 15.0  | 171    | 1  | Q0V2H1 | dual specificity p |
| 27         | 147     | 15.0  | 171    | 1  | A42514 | dual specificity p |
| 28         | 144     | 14.7  | 171    | 1  | I36845 | dual specificity p |
| 29         | 144     | 14.7  | 171    | 2  | T28522 | probable dual spec |

## C:Genetics:

A:Gene: CESP:F26A3.4  
A:Map position: 1  
A:Introns: 117/2; 150/3; 186/3

Query Match 28.0%; Score 275; DB 2; Length 226;

Best Local Similarity 36.8%; Pred. No. 3.3e-19;

Matches 68; Conservative 42; Mismatches 63; Indels 12; Gaps 7;

Qy 10 VQFR-QPSVSGLSQITKSLYSNGVAANKMLSSNQITWTVNVSVEVN-TLYEDIQYM 67

Db 3 LSEFVNPEYAMSEIVPGLFIC-GVSALSKDEMKKHITHIINATTEVPLNRSGLDIQRT 61

Qy 68 QVPVADSPNSRLCDFDFPIADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHMSL 127

Db 62 KLWLEDTPQTYIYPHLEQLSDIQIALIADGGKVLVHCVAGVSRASICLAFLLKYRCNL 121

Qy 128 LDAHTWTKSCRPIIRPNRSGFWEQLHYEFQFGKN--TVHMY---SSVPGMIPDIYEK 182

Db 122 REAYHLMKSRSMVRPNLGFWRQLIAYE-QNVKENAGSVRLVRDEAOPEQLLPDVI---L 177

Qy 183 RLMP 187

Db 178 NIAIP 182

## RESULT 3

S29090

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human

N:Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase, non

C:Species: Homo sapiens (man)

C:Date: 25-Feb-1994 #sequence\_revision 02-May-1994 #text\_change 11-Jun-1999

A:Accession: S29090; A53052

R:Keyse, S.M.; Emelie, E.A.

Nature 359, 644-647, 1992

A:Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine

A:Reference number: S29090; MUID:93024952; PMID:1406996

A:Accession: S29090

A:Molecule type: mRNA

A:Residues: 1-367 <KEY>

A:Cross-references: EMBL:X68277; NID:g29980; PIDN:CAA48338.1; PID:g29981

R:Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.

J. Biol. Chem. 269, 3596-3604, 1994

A:Title: Isolation and characterization of a human dual specificity protein-tyrosine pho

A:Reference number: A53052; MUID:94148864; PMID:8106404

A:Accession: A53052

A:Molecule type: DNA

A:Residues: 1-367 <KWA>

A:Experimental source: leukocyte

A:Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804,

## C:Genetics:

A:Gene: GDB:DUSP1; PTPN10

A:Cross-references: GDB:136197; OMIM:600714

A:Map position: 3q34-5q34

C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specificity

C:Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced

F:181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <VHL>

F:258/Active site: Cys (phosphocysteine intermediate) #status predicted

F:264/Binding site: substrate phosphate (Arg) #status predicted

## Query Match

Best Local Similarity 25.7%; Score 252; DB 1; Length 367;

Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps 1;

Qy 5 SCAPFVQRPQPSVSGLSQITKSLYSNGVAANKMLSSNQITWTVNVSVEVNTLYEDI 64

Db 163 SCSTPLY----DQGGPVEILPFLYLGSAYHASRKDMLDALGITALINVSANCPNHFEGHY 218

Qy 65 QYQMPVPADSPNSRLCDFDFPIADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHA 124

Db 219 QYKSIPIVEDNHKADISSWFNEAIDFISKNAGRVFVHCQAGISRATICLAYLMRTNR 278

Qy 125 MSLDAHTWTKSCRPIIRPNRSGFWEQLHYEFQFGKNTVHMVSSPVGMI 176

Db 279 VKLDEAFEFVKQRSSIIISPNFSFMQLLQFESQVLAPHCSAEAGSPAMAVLD 330

## RESULT 4

S24411

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - mouse

N:Alternate names: 3CH134 protein; protein-tyrosine-phosphatase erp, nonreceptor type

C:Species: Mus musculus (house mouse)

C:Date: 19-Feb-1994 #sequence\_revision 02-May-1994 #text\_change 11-Jun-1999

C:Accession: A54681; S24411

R:Noguchi, T.; Metz, R.; Chen, L.; Mattel, M.G.; Carrasco, D.; Bravo, R.

Mol. Cell. Biol. 13, 5195-5205, 1993

A:Title: Structure, mapping, and expression of erp, a growth factor-inducible gene

A:Reference number: A54681; MUID:93360956; PMID:8355678

A:Accession: A54681

A:Molecule type: DNA

A:Residues: 1-367 <NOG>

A:Cross-references: GB:S64851; NID:g409976; PIDN:AAB27882.1; PID:g409977

R:Charles, C.H.; Abler, A.S.; Lau, L.F.

Oncogene 7, 187-190, 1992

A:Title: cDNA sequence of a growth factor-inducible immediate early gene and characte

A:Reference number: S24411; MUID:92158357; PMID:1741163

A:Accession: S24411

A:Molecule type: mRNA

A:Residues: 1-367 <CHA>

A:Cross-references: EMBL:X61940; NID:g49735; PIDN:CAA43944.1; PID:g49736

C:Genetics:

A:Gene: erp

C:Introns: 123/1; 172/1; 245/1

C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specific

C:Keywords: immediate-early protein; phosphoprotein; phosphoric monoester hydrolase

F:181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <VHL>

F:258/Active site: Cys (phosphocysteine intermediate) #status predicted

F:264/Binding site: substrate phosphate (Arg) #status predicted

## Query Match

Best Local Similarity 25.7%; Score 252; DB 1; Length 367;

Matches 59; Conservative 29; Mismatches 80; Indels 4; Gaps 1;

Qy 5 SCAPFVQRPQPSVSGLSQITKSLYSNGVAANKMLSSNQITWTVNVSVEVNTLYEDI 64

Db 163 SCSTPLY----DQGGPVEILPFLYLGSAYHASRKDMLDALGITALINVSANCPNHFEGHY 218

Qy 65 QYQMPVPADSPNSRLCDFDFPIADHHSVENKQGRITLLHCAAGVSRSAALCLAYLMKYHA 124

Db 219 QYKSIPIVEDNHKADISSWFNEAIDFISKNAGRVFVHCQAGISRATICLAYLMRTNR 278

Qy 125 MSLDAHTWTKSCRPIIRPNRSGFWEQLHYEFQFGKNTVHMVSSPVGMI 176

Db 279 VKLDEAFEFVKQRSSIIISPNFSFMQLLQFESQVLAPHCSAEAGSPAMAVLD 330

## RESULT 5

S52265

dual specificity phosphatase (EC 3.1.3.-) 1 - rat

N:Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase,

C:Species: Rattus norvegicus (Norway rat)

C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 22-Jun-1999

C:Accession: S52265

R:Muda, M.; Schlegel, W.; Arkinstall, S.

Submitted to the EMBL Data Library, January 1995

A:Description: Pathways regulating CL100 gene expression in pituitary cells.

A:Reference number: S52265

A:Accession: S52265

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-367 <MUD>

A:Cross-references: EMBL:X84004; NID:g642264; PIDN:CAA58828.1; PID:g642265

C:Superfamily: dual specificity phosphoprotein phosphatase 1; Vhl-type dual specific

C:Keywords: phosphoprotein; phosphoric monoester hydrolase

F:181-312/Domain: Vhl-type dual specificity phosphoprotein phosphatase homology <VHL>

F:258/Active site: Cys (phosphocysteine intermediate) #status predicted



F:202-333/Domain: VH1-type dual specificity phosphoprotein phosphatase hom  
F:279/Active site: Cys (phosphocysteine intermediate) #status predicted

A;Reference number: Z18439  
A;Accession: T15969

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-619 <CHI>

A:Cross-references: EMBL:U23178; NID:g726421; PID:g726422; PIDN:AAC46719.1; CESP:F08B1.1

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F08B1.1

A:Map position: 57/L; 99/L; 160/3; 252/2; 549/3

Query Match 23.8%; Score 233; DB 2; Length 619;

Best Local Similarity 33.3%; Pred. No. 1.6e-14;

Matches 48; Conservative 35; Mismatches 59; Indels 2; Gaps 1;

Qy 14 QPSVSLGSLQITKSLYSINGVAANNKMLSSNOITWVINVSVVVNT--LYEDIOYMQVPV 71

Db 132 QPTGGIYLTPTNYIGSQIDSLDLDALDISVIVNLMTCPKSVCKEDKNFMRIPV 191

Qy 72 ADSPNSRLCDFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 131

Db 192 NDSYQEKLSPPYPMAYEFLEKCRRAKKCLHCLAGISRPTLAISIMRYMKMGSDDAY 251

Qy 132 TWTKSCRPIIRPNSGFWEOLHYE 155

Db 252 RYVKRRPISPNFNFMGQLLEYE 275

#### RESULT 10

Tl6056

hypothetical protein Fl3D11.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 23-Mar-2001

C:Accession: Tl6056

R:Fulton, L.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid Fl3D11.

A:Reference number: 569020

A:Accession: Tl6056

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-186 <FUI>

A:Cross-references: EMBL:U40939; NID:gl072175; PID:gl072176; PIDN:AAA81700.1; CESP:Fl3D11.3

C:Genetics:

A:Gene: CESP:Fl3D11.3

A:Introns: 30/3; 57/3; 85/3; 125/3; 172/3

Query Match 23.6%; Score 231.5; DB 2; Length 186;

Best Local Similarity 31.4%; Pred. No. 4.3e-15;

Matches 55; Conservative 36; Mismatches 61; Indels 23; Gaps 4;

Qy 17 VSGLSQITKSLY----ISNGVAANNKMLSSNOITWVINVSVVVNTLYEDIOYMQVPVA 72

Db 20 LAGYGCITPSLLKQYNITHGVDCN--LTKTPI-----KGLDRIEVPVD 61

Qy 73 DSPNSRLCDFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLMKYHAMSLLDAH 132

Db 62 DNTLAKITQYEPVVKYEDAKQQGHNVIVYCAAGVSRSATITIVYLMVTENLSLEAYL 121

Qy 133 WTKSCRPIIRPNSGFWEOLHYEFOLFCKNTVHMVSSPVGMI-IPDIYEKVRLEMI 186

Db 122 QVNQRVPIISPNIGFWRQMIDFEKORNGNASVELISGRMARVPVSVLLRRVTII 176

#### RESULT 11

T32494

hypothetical protein C05B10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T32494

R:Geisler, C.; Wamsley, P.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid C05B10.

A:Reference number: 221178

A:Accession: T32494

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-365 <GET>

A:Cross-references: EMBL:AF036585; PIDN:AAB8308.1; GSPDB:GN00022; CESP:C05B10.1

A:Experimental source: strain Bristol N2; clone C05B10

C:Genetics:

A:Gene: CESP:C05B10.1

A:Map position: 4

A:Introns: 22/2; 46/2; 74/3; 163/2; 204/2; 248/1; 287/2

Query Match 22.8%; Score 223.5; DB 2; Length 365;

Best Local Similarity 33.3%; Pred. No. 6.4e-14;

Matches 56; Conservative 25; Mismatches 68; Indels 19; Gaps 4;

Qy 8 FPVQRPQSVSGLSQITKSLYSINGVAANNKMLSSNOITWVINVSVVVNTLYED--IQ 65

Db 180 FPV-----KLTNFLYLGNATKRDVLYKYSISHVINVTSLNPTFEEDPNMR 228

Qy 66 YMQVPVADSPNSRLCDFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLMKYHAM 125

Db 229 YLRISADDNASHNLTKFFPEAISFIDDARRNDSACLHCLAGISRSVTICLAYLMKTEMC 288

Qy 126 SLDDAHTWTKSCRPIIRPNSGFWEOLHYEFOLFCKNTVHMVSSPVGMI 173

Db 289 TLDSAYEVMQVRNASIAPNFHFMGQLTDYE-KMLGLN-----SNRQGV 330

#### RESULT 12

I38890

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human

N:Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvH

C:Species: Homo sapiens (man)

C:Date: 16-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 11-Jun-1999

C:Accession: I38890; A55313

R:Kwak, S.P.; Dixon, J.E.

J. Biol. Chem. 270, 1156-1160, 1995

A:Title: Multiple dual specificity protein tyrosine phosphatases are expressed and re

A:Reference number: A55432; MUID:95138103; PMID:7836374

A:Accession: I38890

A:Molecule type: mRNA

A:Residues: 1-384 <RES>

A:Cross-references: EMBL:U16996; NID:g642012; PIDN:AAB06261.1; PID:g642013

A:Experimental source: placenta

R:Ishibashi, T.; Bottaro, D.P.; Michielli, P.; Kelley, C.A.; Aaronson, S.A.

J. Biol. Chem. 269, 29897-29902, 1994

A:Title: A novel dual specificity phosphatase induced by serum stimulation and heat s

A:Reference number: A55313; MUID:95050849; PMID:7961985

A:Accession: A55313

A:Molecule type: mRNA

A:Residues: 1-8,'GHV',12-70,'R',72-104,'F',107-362,'RCLPTQSSSAAELMQRPNPARTGMESAQPO

A:Cross-references: GB:U15932; NID:g606971; PIDN:AAA64693.1; PID:g606972

A:Experimental source: mammary epithelial cells

C:Genetics:

A:Gene: GDB:DUSP5

A:Cross-references: GDB:385447

A:Map position: 10q25-10q25

C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHI-type dual specific

C:Keywords: phosphoprotein; phosphoric monoester hydrolase

F:186-317/Domain: VHI-type dual specificity phosphoprotein phosphatase homology <VHI>

F:263/Active site: Cys (phosphocysteine intermediate) #status predicted

F:269/Binding site: substrate phosphate (Arg) #status predicted

Query Match 22.5%; Score 221; DB 1; Length 384;

Best Local Similarity 34.3%; Pred. No. 1.2e-13;

Matches 48; Conservative 24; Mismatches 68; Indels 0; Gaps 0;

Qy 19 GLSQITKSLYSINGVAANNKMLSSNOITWVINVSVVVNTLYEDIOYMQVPVADSPNSR 78

Db 178 GPVEILPPLYLGSAYHASKCEFLANLHITALLNVSRRSEACMTLHLYKWIPIVEDSHAD 237

Qy 79 LCDFFDPIADHIHSEVMKQGRTHLHCAAGVSRSAALCLAYLMKYHAMSLLDAHTWTKSCR 138

Db 238 ISSHFQEAIDFIDCVREKGGKVLVHCEAGISRSPTICMAYLMKTKQFLKEAFDIKORR 297

F;263/Binding site: substrate phosphate (Arg) #status predicted

Job time : 29 sec

Search completed: April 21, 2003, 18:34:58  
Job time : 29 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 16:59:44 ; Search time 46 seconds  
(without alignments)  
5560.183 Million cell updates/sec

Title: US-09-527-376-1

Perfect score: 834

Sequence: 1 ctttctctgtatttttgc.....aaactaattgtgagatggtg 834

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 177.6 | 21.3        | 1729   | 4  | US-09-045-973-6   |
| 2          | 87    | 10.4        | 1238   | 2  | US-08-530-290-11  |
| 3          | 81.2  | 9.7         | 1987   | 2  | US-08-990-379-1   |
| 4          | 81.2  | 9.7         | 1993   | 2  | US-08-990-379-2   |
| 5          | 61.6  | 7.4         | 1691   | 3  | US-09-013-881-12  |
| 6          | 58.8  | 7.1         | 789    | 4  | US-09-164-193-6   |
| 7          | 58.8  | 7.1         | 789    | 4  | US-09-221-448A-6  |
| 8          | 58.8  | 7.1         | 1016   | 4  | US-09-164-193-4   |
| 9          | 58.8  | 7.1         | 1016   | 4  | US-09-221-448A-4  |
| 10         | 52.2  | 6.3         | 531    | 4  | US-09-704-139-3   |
| 11         | 52.2  | 6.3         | 594    | 4  | US-09-163-833-3   |
| 12         | 52.2  | 6.3         | 912    | 4  | US-09-163-833-1   |
| 13         | 52.2  | 6.3         | 1390   | 4  | US-09-704-139-1   |
| 14         | 45    | 5.4         | 216    | 2  | US-08-530-290-9   |
| 15         | 38.2  | 4.6         | 970    | 4  | US-09-071-035-183 |
| 16         | 38.2  | 4.6         | 1110   | 4  | US-09-071-035-181 |
| 17         | 37.4  | 4.5         | 1293   | 4  | US-09-078-294-10  |
| 18         | 37    | 4.4         | 7218   | 1  | US-08-232-463-14  |
| 19         | 35.8  | 4.3         | 599    | 4  | US-09-078-294-23  |
| 20         | 34.4  | 4.1         | 289    | 4  | US-09-007-005-17  |
| 21         | 34.4  | 4.1         | 289    | 4  | US-09-244-796-17  |
| 22         | 34.4  | 4.1         | 677    | 2  | US-08-666-405-9   |
| 23         | 34.4  | 4.1         | 882    | 2  | US-08-666-405-27  |
| 24         | 34.4  | 4.1         | 1730   | 3  | US-09-222-817-11  |
| 25         | 34.4  | 4.1         | 1730   | 3  | US-09-222-817-13  |
| 26         | 34.4  | 4.1         | 1730   | 4  | US-09-222-786-11  |
| 27         | 34.4  | 4.1         | 1730   | 4  | US-09-222-786-13  |

Sequence 15, Appl  
Sequence 37, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 134, App  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 18, Appl  
Sequence 18, Appl  
Sequence 611, App  
Patent No. 5447867  
Sequence 2, Appl  
Sequence 116, App

28 34.2 4.1 17327 1 US-07-906-871-15  
29 33.2 4.0 937 3 US-09-248-335-37  
30 32.6 3.9 5798 2 US-08-483-101-1  
31 32.4 3.9 861 1 US-07-988-273-1  
32 32.4 3.9 861 5 PCT-US93-12019-1  
33 32.4 3.9 1011 3 US-09-095-163-1  
34 32.4 3.9 1029 4 US-09-485-648-5  
35 32.4 3.9 1029 4 US-09-503-565-5  
36 32.4 3.9 1029 4 US-09-485-649-5  
37 32.4 3.9 1345 1 US-08-525-654A-134  
38 32.4 3.9 3332 4 US-09-448-806C-1  
39 32.4 3.9 4403765 4 US-09-103-840A-2  
40 32 3.8 1393 1 US-08-174-467-18  
41 32 3.8 1393 3 US-08-452-071-18  
42 31.8 3.8 1023 4 US-09-134-001C-611  
43 31.8 3.8 1636 6 5447867-2  
44 31.8 3.8 2235 4 US-09-153-804-2  
45 31.8 3.8 3079 4 US-09-643-597-116

#### ALIGNMENTS

RESULT 1  
US-09-045-973-6  
; Sequence 6, Application US/09045973  
; Patent No. 6165767  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Baughn, Mariah  
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/045,973  
; FILING DATE: Filed Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0491 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1729 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: BRSTNOT16  
; CLONE: 3041794  
US-09-045-973-6



|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |      |                                                                 |      |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|-----------------------------------------------------------------|------|
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 206  | CACCATGGTCATCAATGCTCTCACTGGAGGTAGTGAACACCTTGTTATGAGGATATCCAGTA  | 265  |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1046 | CACGGCTCTACTGAATGCTCTCTCAGACTGCCCCCAATCACTTTTGAGGGACATTTACCAGTA | 1101 |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 266  | CATGCAGGTACCTGTGGCTGACTCCCTAACTCACGCTCTCTGTGACTTCTTTTGAACCTAT   | 325  |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1106 | CAAGTGCATCCGGTAGAAGATAACCAAGGCTGCATCATGCTCTGTTTCATGGAAGC        | 1165 |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 326  | TGCTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCGTACTTTTCTGCATCTGTCTGC    | 385  |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1166 | CATCGAATACATAGACGCACTGAAGGACTGCCGAGGGCGAGTCTGTTTCTGCTGCGCAGC    | 1225 |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 386  | TGGTGTGAGCCCTCAGCTGCGCTGTGCGCTACCTCATGAAGTATACACGCCATGTC        | 445  |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1226 | CGGCATCTCTAGATCAGCCACCATCTGCTGGCTTACCTGATGATGAAGAAACGGGTGAG     | 1285 |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 446  | CTGCTGACGCGCCACACGCTGGACCAAGTCATCCGCGCCCATCATCCGAGCCCAACAGCGG   | 505  |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1286 | GCTGAGGAGGCTTTGAGTTCGTCGAAGCAGGCGCGCTAGCATCATCTCGCCCAACTTCAG    | 1345 |
| QY                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 506  | CTTTTGGGAGCAGTCATCCATGATGAGTTCCTCAATTTGTTGGCAAG                 | 551  |
| DB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1346 | CTTCATGGCCAGTTCCTGCAGTTCGAGTTCAGGTGCTCAGGCACG                   | 1391 |
| <p>RESULT 5</p> <p>US-09-013-881-12</p> <p>; Sequence 12, Application US/09013881</p> <p>; Patent No. 6132964</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Bandman, Olga</p> <p>; APPLICANT: Lal, Preeti</p> <p>; APPLICANT: Hillman, Jennifer L.</p> <p>; APPLICANT: Corley, Neil C.</p> <p>; APPLICANT: Guegler, Karl J.</p> <p>; APPLICANT: Shah, Purvi</p> <p>; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES</p> <p>; NUMBER OF SEQUENCES: 16</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: Incyte Pharmaceuticals, Inc.</p> <p>; STREET: 3174 Porter Drive</p> <p>; CITY: Palo Alto</p> <p>; STATE: CA</p> <p>; COUNTRY: USA</p> <p>; ZIP: 94304</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Diskette</p> <p>; COMPUTER: IBM Compatible</p> <p>; OPERATING SYSTEM: DOS</p> <p>; SOFTWARE: FastSeq for Windows Version 2.0</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/09/013,881</p> <p>; FILING DATE: HEREWITH</p> <p>; CLASSIFICATION:</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER:</p> <p>; FILING DATE:</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: BILLINGS, LUCY J.</p> <p>; REGISTRATION NUMBER: 36,749</p> <p>; REFERENCE/DOCKET NUMBER: PF-0470 US</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: 650-855-0555</p> <p>; TELEFAX: 650-845-4166</p> <p>; TELEX:</p> <p>; INFORMATION FOR SEQ ID NO: 12:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 1691 base pairs</p> <p>; TYPE: nucleic acid</p> <p>; STRANDEDNESS: single</p> <p>; TOPOLOGY: linear</p> <p>; IMMEDIATE SOURCE:</p> <p>; LIBRARY: MUSCNOT02</p> |      |                                                                 |      |





```
RESULT 9
US-09-221-448A-4
; Sequence 4, Application US/09221448A
; Patent No. 6436685
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSATTP PROTEIN MOLECULES AND USES THEREFOR (AS AMENDED)
; FILE REFERENCE: MNI-051DV1
; CURRENT APPLICATION NUMBER: US/09/221.448A
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1016
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(789)
US-09-221-448A-4

Query Match          7.18; Score 58.8; DB 4; Length 1016;
Best Local Similarity 57.28; Pred. No. 3.7e-08;
Matches 127; Conservative 0; Mismatches 92; Indels 3; Gaps 1;

Qy 254 GGATATCCAGTACATGCGAGGTACCTGTGGCTGACTCCCTTAACCTACAGCTCTGTGACTT 313
Db 198 GGGATCCGCTACCTGGTGTGTAGGCCCCACGACTCGCCAGCCTTTGACATGAGCATCCA 257

Qy 314 CTTTGACCTATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCG---TACTTTT 370
Db 258 CTTCCAGACGGCTCGCGACTTCATCCACCGGCGCTGAGCCAGCAGGAGGAAGATCCT 317

Qy 371 GCTGCACTGCTGCTGTGTGTGAGCGCTGAGCTGCCCTGTGCTGCTGCTGCTGCTGCTG 430
Db 318 GGTGCACTGCTGCTGTGCGGTGAGCGGATCGCCACCTGCTGCTGCTGCTGCTGCTGCTG 377

Qy 431 GTACCACCCATGTCCTCTGCTGGAGCGCCACAGCTGGACCAA 472
Db 378 GTACCACCCATGTCCTCTGCTGGAGCGCCATCAAGNAAGTCAA 419

RESULT 10
US-09-704-139-3
; Sequence 3, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704.139
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185.772
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-704-139-3

Query Match          6.3%; Score 52.2; DB 4; Length 531;
Best Local Similarity 44.8%; Pred. No. 2.7e-06;
Matches 201; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

Qy 147 CTGTATATCAGCAATGTTGTGCGCCCAACACAAAGCTCATGCTGCTGAACACAGATC 206
Db 76 CTCCTTCTCGGAGCGCGGAGCGCGCGGCGGAGGAGCAGCTGGCGCGCGGGAGTC 135

Qy 207 ACCATGGTATCAATGTTCTAGTGGAGGTAGTGAACACCTTTGATGAGGATATCCAGTAC 266
Db 136 ACGCTGTGCTCAACGTTCTCCCGCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGAG 195
```

```
Qy 267 ATGCAGGTACCTGTGGCTGACTCCCTAACTACAGTCTCTGTGACTTCTTTGACCTATT 326
Db 196 CTGCGCGTGGCCGTTTCGAGGACGCCGGCTGAGGACCTGTGGCGCAGCTGGAGCCACG 255

Qy 327 GCTGACCATATCCACAGCGTGGAGATGAACGAGGCGCCTACTTTGCTGCACTGTGCTGCT 386
Db 256 TGCGCCGCATGAGGCGCGGTGCGCCCGCGCGGCTGCCCTAGTCTACTTGCAAGAAC 315

Qy 387 GGTGTGAGCCGCTCAGCTGCGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446
Db 316 GCGCGCAGCGCTGCGCGCGCTGCTGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375

Qy 447 CTGCTGGAGCGCCACAGCTGAGCAAGTCAATCGCGGCCCATCATCCGACCCCAACAGCGC 506
Db 376 CTGGGAAGGCCCTTCCAGATGTTGAAGAGCGCTGCGCCCGCTAGCAGAACCCCGGGC 435

Qy 507 TTTTGGGAGCAGCTCATCCACTATGAGTTCCAATTTGTTGGCAAGAACACTGTGACACATG 566
Db 436 TTTGTTCTCAGCTCCAGAGTATGAGGAGGCCCTCCAGGCGCCCTGCTGCTGCTGCTGCTG 495

Qy 567 GTCAGTTCCTCCAGTGGGAATGATCCCTGA 595
Db 496 GAGCCCCCAGCCTTAGGTTGGGCCCTGA 524

RESULT 11
US-09-163-833-3
; Sequence 3, Application US/09163833
; Patent No. 6268135
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: No. 6268135el Phospholipase Molecule and Uses Therefor
; FILE REFERENCE: mni-059
; CURRENT APPLICATION NUMBER: US/09/163.833
; CURRENT FILING DATE: 1998-09-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(594)
US-09-163-833-3

Query Match          6.3%; Score 52.2; DB 4; Length 594;
Best Local Similarity 53.9%; Pred. No. 2.8e-06;
Matches 130; Conservative 0; Mismatches 108; Indels 3; Gaps 1;
```

```
Qy 279 GTGGCTGACTCCCTAACTCAGTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATC 338
Db 310 GCGGACGACAAACCCCTTCTCGACCTCAGTGTCTACTTTGCTGCTGCTGCTGCTGCTG 369

Qy 339 C---ACAGCGTGGAGATGAAGCAGGCGCTACTTTGCTGCACTGTGCTGCTGCTGCTGAGC 395
Db 370 CGAGCTGCCCTCAGTGTTCCTCCAGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429

Qy 396 CGCTCAGCTGCGCTGTGCTGCGCTTACCTATCAAGTACCAGCCATGCTCCCTGCTGGAC 455
Db 430 CGCTCTGCCACACTGTCTTCCCTGCTGCTTCTCATGATCTATGAGAACATGACGCTGCTAG 489

Qy 456 GCCACAGCTGAGCAAGTCAATGCGCGCCCATCATCCGACCCCAACAGCGGCTTTTGGAG 515
Db 490 GCCATCCAGAGGTGAGGCGCCACCGCAATATCTGCCCTAACTCAGGCTTCTCCCGGAG 549

Qy 516 C 516
Db 550 C 550

RESULT 12
```

US-09-163-833-1  
; Sequence 1, Application US/09163833  
; Patent No. 6268135  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NO. 6268135el Phospholipase Molecule and Uses Therefor  
; FILE REFERENCE: mni-059  
; CURRENT APPLICATION NUMBER: US/09/163,833  
; CURRENT FILING DATE: 1998-09-30  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 912  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (52)..(645)  
US-09-163-833-1  
Query Match 6.3%; Score 52.2; DB 4; Length 912;  
Best Local Similarity 53.9%; Pred. No. 3.6e-06;  
Matches 130; Conservative 0; Mismatches 108; Indels 3; Gaps 1;  
QY 279 GTGCTGACCTCCCTAACTCAGCTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATC 338  
DB 361 GCGGACGACACCCCTTCTTCGACCTCAGTGTCTACTTCTGCTGTGCTCGATATC 420  
QY 339 C---ACAGCTGGAGTAGAAGAGGCGCGTACTTGTGCTGACACTGTGCTGTGCTGAGC 395  
DB 421 CGAGCTGCCCTCAGTGTTCGCCAAGGCGCGTGTGTGTACACTGTGCTCATGCGGTAAAGC 480  
QY 396 CGCTCAGCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 455  
DB 481 CGCTCTGCCACACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 456 GCCACACGCTGGACCAAGTATGCGGCCCATCATCGACCCCAACAGCGGCTTTTGGGAG 515  
DB 541 GCCATCCAGAGGTGACGAGCGCCAGCGCAATATCTGCCTAACTCAGGCTTCTCCGCGAG 600  
QY 516 C 516  
DB 601 C 601  
RESULT 13  
US-09-704-139-1  
; Sequence 1, Application US/09704139  
; Patent No. 6420153  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller, Rosanna  
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR  
; FILE REFERENCE: 10448-018001  
; CURRENT APPLICATION NUMBER: US/09/704,139  
; CURRENT FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: US 60/185,772  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 1390  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (329)..(859)  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1390)  
; OTHER INFORMATION: n = a, t, g, or c  
US-09-704-139-1  
Query Match 6.3%; Score 52.2; DB 4; Length 1390;  
Best Local Similarity 44.8%; Pred. No. 4.6e-06;

Matches 201; Conservative 0; Mismatches 248; Indels 0; Gaps 0;  
QY 147 CTGTATATCAGCAATGTGTGGCGCAACAACAAGCTCATGTCTAGCAACACAGATC 206  
DB 404 CTCTTCTCGGAGCGCGGCGGCGGAGGAGGAGTGGCGCGCGGGGAGTC 463  
QY 207 ACCATGTCATCAATGTCTCAGTGGAGTAGTGAACACCTTGTATGAGGATATCCAGTAC 266  
DB 464 ACCTGTGCGTCAAGCTCTCCCGCCAGCAGCCCGCGCGCGCGCGCGGTGGCAGAG 523  
QY 267 ATCAGGTACTGTGGCTGACTCCCTAACTCAGTCTCTGTGACTTCTTTGACCCCTATT 326  
DB 524 CTGCGGTGCGCGTGTTCGACGACCCCGGCTGAGGACCTGCTGGCGCACCTGGAGCCACG 583  
QY 327 GCTGACCATATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTTGTGCACTGTGCTGCT 386  
DB 584 TGGCGCCCATGAGGCGCGGTGGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 643  
QY 387 GGTGTGAGCGCTCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446  
DB 644 GCGCGCAGCGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCGCTG 703  
QY 447 CTGCTGAGCGCCACAGCTGAGCAAGTATGCGCGCGCGCTGCTGCGCGCGCTGCTGCGCG 506  
DB 704 CTGCGCAAGCGCTTCCAGATGTTGAAGAGCGCTGCGCGCGCTGAGCAGAAACCGCGGC 763  
QY 507 TTTTGGGAGCAGCTCATCCACTATGATTTCCATTGTTTGGCAAGAACACTGTGCGCATG 566  
DB 764 TTTGCTGCTCAGCTGCGAAGTATGAGAGGCGCTTCCAGCGCGCGCTGCTGCTGCTGCTG 823  
QY 567 GTCAGTTCCCGCAGTGAATGATCCCTGTA 595  
DB 824 GAGCCCCCAGCCTTAGGGTTGGCGCTGA 852  
RESULT 14  
US-08-530-290-9  
; Sequence 9, Application US/08530290  
; Patent No. 5958721  
; GENERAL INFORMATION:  
; APPLICANT: Marshall, Christopher John  
; APPLICANT: Ashworth, Alan  
; APPLICANT: Hughes, David Anthony  
; TITLE OF INVENTION: Methods for Screening of Substances for  
; TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/530,290  
; FILING DATE: 14-DEC-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/GB94/00694  
; FILING DATE: 31-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9402573.1  
; FILING DATE: 10-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9307250.2  
; FILING DATE: 07-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 16:45:49 ; Search time 1070 Seconds

(without alignments)  
12623.398 Million cell updates/sec

Title: US-09-527-376-1

Perfect score: 834

Sequence: 1 ctttctctgtattttttgc.....aaactaattgtgagatggtg 834

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estopl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: gb\_gss.\*

18: em\_gss\_hum.\*

19: em\_gss\_inv.\*

20: em\_gss\_pln.\*

21: em\_gss\_vrt.\*

22: em\_gss\_fun.\*

23: em\_gss\_mam.\*

24: em\_gss\_mus.\*

25: em\_gss\_other.\*

26: em\_gss\_pro.\*

27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description        |
|------------|-------|---------------|--------|----|--------------------|
| 1          | 611   | 73.3          | 611    | 10 | BE222374 hul1d05.x |
| 2          | 543.6 | 65.2          | 592    | 9  | AA527292 ng39g09.s |
| 3          | 541.2 | 64.9          | 547    | 9  | AA915932 on18c06.s |
| 4          | 495.8 | 59.4          | 746    | 10 | BE563213 601335506 |
| 5          | 490.2 | 58.8          | 495    | 9  | AI215158 qp41h06.x |
| 6          | 464.6 | 55.7          | 489    | 9  | AA926744 om25a02.s |

|   |    |       |      |      |    |          |
|---|----|-------|------|------|----|----------|
| C | 7  | 441   | 52.9 | 441  | 9  | AI283262 |
|   | 8  | 423.6 | 50.8 | 1161 | 11 | AK015917 |
|   | 9  | 413.8 | 49.6 | 446  | 13 | BI037896 |
|   | 10 | 359.8 | 43.1 | 894  | 14 | EQ934977 |
|   | 11 | 356.4 | 42.7 | 679  | 12 | EG333140 |
| C | 12 | 341.8 | 41.0 | 457  | 10 | AW972251 |
|   | 13 | 324.6 | 38.9 | 744  | 13 | BI907369 |
|   | 14 | 313.4 | 37.6 | 840  | 11 | AK007061 |
|   | 15 | 305.2 | 36.3 | 784  | 13 | BI914744 |
|   | 16 | 302.6 | 36.3 | 625  | 10 | BB661786 |
|   | 17 | 291.8 | 35.0 | 429  | 12 | BF559942 |
|   | 18 | 290.2 | 34.8 | 674  | 12 | EG423899 |
|   | 19 | 286.8 | 34.4 | 813  | 12 | EG619206 |
| C | 20 | 283.4 | 34.0 | 299  | 9  | AI916043 |
|   | 21 | 280.2 | 33.6 | 292  | 9  | AA356476 |
|   | 22 | 276.4 | 33.1 | 1102 | 13 | BM550612 |
|   | 23 | 269.2 | 32.3 | 563  | 10 | AW844028 |
| C | 24 | 229   | 27.5 | 234  | 9  | AI470481 |
|   | 25 | 226   | 27.1 | 562  | 10 | AW843834 |
| C | 26 | 217.8 | 26.1 | 493  | 9  | AI025489 |
|   | 27 | 201.4 | 24.1 | 491  | 9  | AA813372 |
| C | 28 | 197.4 | 23.7 | 925  | 12 | BF160291 |
|   | 29 | 191   | 22.9 | 818  | 13 | BI464358 |
|   | 30 | 190.6 | 22.9 | 571  | 9  | AL711163 |
|   | 31 | 187.2 | 22.4 | 1089 | 11 | AK009744 |
|   | 32 | 185.4 | 22.2 | 826  | 13 | BI556609 |
| C | 33 | 184.2 | 22.1 | 460  | 9  | AA725688 |
|   | 34 | 183   | 21.9 | 1021 | 14 | BQ951245 |
| C | 35 | 182.8 | 21.9 | 354  | 12 | BF471540 |
|   | 36 | 181   | 21.7 | 649  | 10 | BB589466 |
|   | 37 | 180.4 | 21.6 | 564  | 10 | BE235660 |
|   | 38 | 177.2 | 21.2 | 1016 | 9  | AL527826 |
|   | 39 | 175.4 | 21.0 | 876  | 12 | BF130916 |
|   | 40 | 175.2 | 21.0 | 1095 | 14 | BM928368 |
|   | 41 | 175   | 21.0 | 884  | 13 | BI768742 |
| C | 42 | 173.8 | 20.8 | 188  | 10 | AW372085 |
|   | 43 | 173.4 | 20.8 | 949  | 14 | BQ939586 |
|   | 44 | 171.4 | 20.6 | 634  | 12 | BG765121 |
| C | 45 | 171.2 | 20.5 | 465  | 9  | AA813123 |

## ALIGNMENTS

RESULT 1  
BE222374/c  
LOCUS hul1d05.xl NCI-CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3166281 3'  
DEFINITION similar to TR:O95147 O95147 MKP-1 LIKE PROTEIN TYROSINE PHOSPHATASE  
; mRNA sequence.  
ACCESSION BE222374  
VERSION BE222374  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 611)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov  
Seq primer: -40UP from Gibco

```
FEATURES
source
High quality sequence stop: 485.
Location/Qualifiers
1. .611
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3166281"
/clone_lib="NCI_CGAP_Lu24"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI_CGAP_Lu24 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT: 151 a 140 c 172 g 148 t
ORIGIN
Query Match 73.3%; Score 611; DB 10; Length 611;
Best Local Similarity 100.0%; Pred. No. 1.5e-158;
Matches 611; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 126 CTCTCGCAGATAACCAAAAGCCTGTATATCAGCAATGGTGTGCGCGCCAAACAAGCTC 185
Db 611 CTCTCGCAGATAACCAAAAGCCTGTATATCAGCAATGGTGTGCGCGCCAAACAAGCTC 552
QY 186 ATGCTGTCTAGCAACAGATACCATGGTTCATCAATGTCTCAGTGGAGGTAGTGAACACC 245
Db 551 ATGCTGTCTAGCAACAGATACCATGGTTCATCAATGTCTCAGTGGAGGTAGTGAACACC 492
QY 246 TTGTATCAGGATATCCAGTACATGCAGGTACCTGTGTGCTGACCTCCCTAACTCAGCTCTC 305
Db 491 TTGTATCAGGATATCCAGTACATGCAGGTACCTGTGTGCTGACCTCCCTAACTCAGCTCTC 432
QY 306 TGTGACTTCTTTGACCCATTGTGTGACCATATCCACAGCTGGAGATGAAGCAGGGCCGT 365
Db 431 TGTGACTTCTTTGACCCATTGTGTGACCATATCCACAGCTGGAGATGAAGCAGGGCCGT 372
QY 366 ACITTTGCTGACGT 425
Db 371 ACITTTGCTGACGT 312
QY 426 ATGAAGTACCAAGCCATGTCCTGCTGGAGCCCAACAGCTGGACCAAGTTCATGCGGCC 485
Db 311 ATGAAGTACCAAGCCATGTCCTGCTGGAGCCCAACAGCTGGACCAAGTTCATGCGGCC 252
QY 486 ATCATCGACCCCAACAGCGGCTTTTGGAGCAGCTCATCCACTATGAGTTTCCAAATTTT 545
Db 251 ATCATCGACCCCAACAGCGGCTTTTGGAGCAGCTCATCCACTATGAGTTTCCAAATTTT 192
QY 546 GGCAAGAACACTGTGCACATGGTTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAG 605
Db 191 GGCAAGAACACTGTGCACATGGTTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAG 132
QY 606 AAGGAAGTCCGTTTGATGATTCACCTGTAGCCATCCACAGCGCCCTGATTTGGAGTCA 665
Db 131 AAGGAAGTCCGTTTGATGATTCACCTGTAGCCATCCACAGCGCCCTGATTTGGAGTCA 72
QY 666 GAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAACTTGAACATTTCTACTTTTGT 725
Db 71 GAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAACTTGAACATTTCTACTTTTGT 12
QY 726 TGATACAGAAA 736
Db 11 TGATACAGAAA 1
RESULT 2
AA527292/c 592 bp mRNA linear EST 05-AUG-1997
LOCUS
```

```

|||||
Db 293 TCCCCAGTGGGAATGATCCCTGACATCTATGAGAGGAAGTCCGTTTGATGATTCACATG 234
QY 633 TGAGCCATCCCGAGGAGCCCTGCATGGAGTGAGGTCAGAGGTACAGATCTATTGTTGATCTTAC 692
Db 233 TGAGCCATCCCGAGGAGCCCTGCATGGAGTCAGAGGTACAGATCTATTGTTGATCTTAC 174
QY 693 ACCAAGATCCAAACTTGAACATCTACTTTTGTGTATACAGAAAAAAGAGATGATGCCT 752
Db 173 ACCAAGATCCAAACTTGAACATCTACTTTTGTGTATACAGAAAAAAGAGATGATGCCT 114
QY 753 TTTATGAGCACAAAAAAGAGTGTCTGCTAGCTTTTAACTTTTATAATCCATTTTTTTTTCAGA 812
Db 113 TTTATGAGCACAAAAAAGAGTGTCTGCTAGCTTTTAACTTTTATAATCCATTTTTTTTTCAGA 54
QY 813 TTAACTAATTTGTGAGATGGTG 834
Db 53 TTAACATAATTTGTGAGATGGTG 32

RESULT 3
AA915932/c
LOCUS
DEFINITION
AA915932 547 bp mRNA linear EST 23-JUN-1998
similar to TR:Q91790 Q91790 MAP KINASE PHOSPHATASE ;, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 547)
NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 834 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 451.
Location/Qualifiers
i. .547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1557034"
/clone_lib="NCI-CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
neuroendocrine lung carcinoid, and was then primed with a
Not I - oligo(dT) primer. Double-stranded cDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 142 a 126 c 153 g 126 t
ORIGIN

Query Match 64.9%; Score 541.2; DB 9; Length 547;
Best Local Similarity 99.5%; Pred. No. 3.5e-139;
Matches 543; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 187 TCGTCTTAGCAACACAGATCACCATGGTTCATCAATGTCAGTGGAGTAGTGAACACCT 246
|||||
Db 546 TCGTCTTAGCAACACAGATCACCATGGTTCATCAATGTCAGTGGAGTAGTGAACACCT 487
QY 247 TGTATGAGGATATCCAGGTACATGCAGGTACCTGTGTGCTGACTCCCTAACTACAGTCTCT 306
|||||
Db 486 TGTATGAGAATATCCAGGTACATGCAGGTACCTGTGTGCTGACTCCCTAACTACAGTCTCT 427
QY 307 GTGACTTCTTTGACCCCTATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTA 356
|||||
Db 426 GTGACTTCTTTGACCCCTATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTA 367
QY 367 CTTTGTGTCACATGCTGCTGTGAGCGGCTCAGCTGCCCTGCTGCTGCTGCTGCTGCTGCT 426
|||||
Db 366 CTTTGTGTCACATGCTGCTGTGAGCGGCTCAGCTGCCCTGCTGCTGCTGCTGCTGCTGCT 307
QY 427 TGAAGTACCACGCCCATGTCCCTGCTGGAGCGCCACACAGTGGACCAAGTCAATGCCGCCCA 486
|||||
Db 306 TGAAGTACCACGCCCATGTCCCTGCTGGAGCGCCATCAGTGGACCAAGTCAATGCCGCCCA 247
QY 487 TCATCCGACCCCAACAGCGGCTTTTGGAGCAGCTCATCCACTATGAGTTCCCAATTTGTTG 546
|||||
Db 246 TCATCCGACCCCAACAGCGGCTTTTGGAGCAGCTCATCCACTATGAGTTCCCAATTTGTTG 187
QY 547 GCAAGAACACTGTGCACATGGTTCAGTTCCCTGAGTGGGAATGATCCCTGACATCTATGAGA 606
|||||
Db 186 GCAAGAACACTGTGCACATGGTTCAGTTCCCTGAGTGGGAATGATCCCTGACATCTATGAGA 127
QY 607 AGGAAGTCCGTTTGATGATTCACATGTGAGCCATCCACAGAGCCCTGCAATGGAGTCAG 666
|||||
Db 126 AGGAAGTCCGTTTGATGATTCACATGTGAGCCATCCACAGAGCCCTGCAATGGAGTCAG 67
QY 667 AGTACAGATCTATTGTTGATCTTACCAAGATCCAAAGTTCGAATCTTACTTTGTT 726
Db 66 AGTACAGATCTATTGTTGATCTTACCAAGATCCAAAGTTCGAATCTTACTTTGTT 7
QY 727 GATACA 732
Db 6 GATACA 1

RESULT 4
BE563213 746 bp mRNA linear EST 15-AUG-2000
LOCUS
DEFINITION
601335506F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3689593 5',
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 746)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: ATCC
CDNA Library Arrayed by: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM384 row: k column: 02
High quality sequence stop: 700.
Location/Qualifiers
i. .746
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

```

```

/clonelib="IMAGE:3689593"
/clonelib="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 194 a 184 c 186 g 182 t
ORIGIN
Query Match 59.4%; Score 495.8; DB 10; Length 746;
Best Local Similarity 97.1%; Pred. No. 1.4e-126;
Matches 611; Conservative 0; Mismatches 7; Indels 11; Gaps 10;
Qy 207 ACCATGGTCATCATGTCTCAGTGGAGTAGTGAACACCTTTGTATGAGGATATCCAGTAC 266
Db 1 ACCATGGTCATCAA-GTCTAGTGGAGTAGTGAACACC-TGTATGAGGATATCCAGTAC 58
Qy 267 ATGAGGTACCTGTGGCTGACTCCCTTAACCTACGTCCTCTGTGACTCTTTGACCCCTATT 326
Db 59 ATGAGGTACCTGTGGCTGACTCCCTTAACCTACGTCCTCTGTGACTCTTTGACCCCTATT 116
Qy 327 GCTGACCATATCCACAGC-GTGGAGATGAAGCAGGCGGTACTTTGTGCACTGTGCTGC 385
Db 117 GCTGACCATATCCACAGCGGTGGAGATGAAGCAGGCGGTACTTTGTGCACTGTGCTGC 176
Qy 386 TGGTGTGAGCGCTCAGCTGCCCTGTGCTGCTTACCTATGATGAAGTACACGCGCATGTC 445
Db 177 TGGTGTGAGCGCTCAGCTGCCCTGTGCTGCTTACCTATGATGAAGTACACGCGCATGTC 236
Qy 446 CTGCTGAGCGCCACACGCTGGACCAAGTCATGCGCGGCCCATCATCGACCCCAACAGCGG 505
Db 237 CTGCTGAGCGCCACACGCTGGACCAAGTCATGCGCGGCCCATCATCGACCCCAACAGCGG 296
Qy 506 CTTTGTGGAGCAGCTCATCCACTATGAGTTCCTCAATTTGTGCAAGACACTGTGCACAT 565
Db 297 C-TTGTGGAGCAGCTCATCCACTATGAGTTCCTCAATTTGTGCAAGACACTGTGCACAT 355
Qy 566 GGTGAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGATGAT 625
Db 356 GGTGAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCG-TTGATGAT 414
Qy 626 TCCACTGTGAGCCATCCACAGCGCCCTGCAATGAGTFCAGAGTACAGATCTATTGTTG 685
Db 415 TCCACTGTGAGCCATCCACAGCGCCCTGCA-TGAGTFCAGAGTACAGATCTA-TGTTG 472
Qy 686 ATCTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGATACAGAAAAAACAGAT 745
Db 473 ATCTTACACCAAGATCCAAACTTGAACATTTCTAC-TTGTGTGATACAGAAAAAACAGAT 531
Qy 746 GATGCCCTTTATGAGCAAAAAAGAGTGTGTGAGCTTTTAACTTTTATATATCCATTTT 805
Db 532 GATGCCCTTTATGAGCAAAAAAGAG-TGCTGTAGCTTTTAACTTTTATATATCCATTTT 590
Qy 806 TTTGAGATTAACATTAATTTGTGAGATGGTG 834
Db 591 TTTGAGATTAACATTAATTTGTGAGATGGTG 619
RESULT 5
AI215158/c AI215158 495 bp mRNA linear EST 02-FEB-1999
LOCUS qp411h06.x1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1925627 3'
DEFINITION similar to WP:F13D11.3 CE04389 PROTEIN TYROSINE-PHOSPHATASE FAMILY
; mRNA sequence.
ACCESSION AI215158
VERSION AI215158.1 GI:3778759
KEYWORDS EST.

```

```

SOURCE ORGANISM human..
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 495)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@nci.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1305 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 444.
FEATURES
Location/Qualifiers
source 1..495
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1925627"
/clone_lib="NCI_CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT 125 a 113 c 138 g 119 t
ORIGIN
Query Match 58.8%; Score 490.2; DB 9; Length 495;
Best Local Similarity 99.4%; Pred. No. 4.9e-125;
Matches 492; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 273 GTACCTGTGGCTGACTCCCTTAACCTACGTCCTGTGACTCTTTGACCCCTATTGCTGAC 332
Db 495 GTACCTGTGGCTGACTCCCTTAACCTACGTCCTGTGACTCTTTGACCCCTATTGCTGAC 436
Qy 333 CATATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGTGCACTGTGCTGGTGTG 392
Db 435 CATATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGTGCACTGTGCTGGTGTG 376
Qy 393 AGCGCTCAGCTGCCCTGTGCTGCCTACCTCATGAAGTACACGCGCATGCTCCCTGCTG 452
Db 375 AGCGCTCAGCTGCCCTGTGCTGCCTACCTCATGAAGTACACGCGCATGCTCCCTGCTG 316
Qy 453 GACGCCACAGGTGACCAAGTCAATGTCGCGCCCATCATCGACCCCAACAGCGGCTTTGG 512
Db 315 GACGCCACAGGTGACCAAGTCAATGTCGCGCCCATCATCGACCCCAACAGCGGCTTTGG 256
Qy 513 GAGCAGCTCATCCACTATGAGTTCCTCAATTTTGGCAAGAACACTGTGCACATGCTCAGT 572
Db 255 GAGCAGCTCATCCACTATGAGTTCCTCAATTTTGGCAAGAACACTGTGCACATGCTCAGT 196
Qy 573 TCCCCAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGATGATTCACATG 632
Db 195 TCCCCAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGATGATTCACATG 136
Qy 633 TGAGCCATCCACAGCGCCCTGCAATGAGTACAGAGTACAGATCTATTGTTGATCTTAC 692
Db 135 TGAGCCATCCACAGCGCCCTGCAATGAGTACAGAGTACAGATCTATTGTTGATCTTAC 76

```



```

QY 693 ACCAAGTCCAACTTGAACATCTACTTTTGTGTATACAGAAAAACAGATGATGCCT 752
      |||||||
Db 75  ACCAAGTCCAACTTGAACATCTACTTTTGTGTATACAGAAAAACAGATGATGCCT 16
      |||||||
QY 753 TTTATGAGCACAAA 767
      |||||||
Db 15  TTTATGAGCACAAA 1
      |||||||

RESULT 6
AA926744/c
LOCUS
DEFINITION
  Am25a02.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
  IMAGE:1542026 3' similar to WP:F26A3.4 CE09669 PROTEIN-TYROSINE
  PHOSPHATASE ;, mRNA sequence.
ACCESSION
  AA926744
VERSION
  AA926744.1 GI:3075641
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 489)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs@mail.nih.gov
  This clone is available royalty-free through LNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 1301 Std Error: 0.00
  Seq primer: -40m13 fwd. Et from Amersham
  High quality sequence stop: 394.
  Location/Qualifiers
    1..489
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1542026"
    /clone_lib="Soares_NFL_T_GBC_S1"
    /lab_host="DH10B"
    /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
    a modified polylinker; Site_1: Not I; Site_2: Eco RI;
    Equal amounts of plasmid DNA from three normalized
    libraries (fetal lung NbHL19w, testis NHT, and B-cell
    NCI-CGAP_GCB1) were mixed, and ss circles were made in
    vitro. Following RAP purification, this DNA was used as
    tracer in a subtractive hybridization reaction. The driver
    was PCR-amplified cDNAs from pools of 5,000 clones made
    from the same 3 libraries. The pools consisted of
    I.M.A.G.E. clones 297480-302087, 682632-687239,
    726408-728711, and 729096-731399. Subtraction by Bento
    Soares and M. Fatima Bonaldo."
  135 a 108 c 123 g 123 t

BASE COUNT
ORIGIN
  Query Match 55.7%; Score 464.6; DB 9; Length 489;
  Best Local Similarity 99.0%; Pred. No. 6.2e-118;
  Matches 478; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 352 TGAAGCAGGGCCGTACTTTGCTGCACGTGCTGCTGTGTGAGCCGCTCAGCTGCCTGT 411
      |||||||
Db 482 TTAAGCAGGGCCGTACTTTGCTGCACGTGCTGCTGTGTGAGCCGCTCAGCTGCCTGT 423
      |||||||
QY 412 GCCTCGCTACCTCATGAGTACCGCATGTCCTGCTGGACGCCACACCTGGACCA 471
      |||||||
Db 422 GCCTCGCTACCTCATGAGTACCGCATGTCCTGCTGGACGCCACACCTGGACCA 364
      |||||||
QY 472 AGTCATGCGCGCCATCATCCGACCAACAGCGGCTTTTGGGAGCAGCTCATCCATATG 531
      |||||||
Db 363 AGTCATGCGCGCCATCATCCGACCAACAGCGGCTTTTGGGAGCAGCTCATCCATATG 304
      |||||||
QY 532 AGTTCCAATTTGTTGGCAAGAACACTGTGCACATGCTGCCAGTGGGAATGATCC 591
      |||||||

```

```

Db 303 AGTTCCAATTTGTTGGCAAGAACACTGTGCACATGTCAGTCCCGAGTGGGAATGATCC 244
      |||||||
QY 592 CTGACATCTATGAGAAGGAGTCGCTTTGATGATTCACATGTCAGCATCCACGAGCCC 651
      |||||||
Db 243 CTGACATCTATGAGAAGGAGTCGCTTTGATGATTCACATGTCAGCATCCACGAGCCC 184
      |||||||
QY 652 CTGCATTGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAACTTGAA 711
      |||||||
Db 183 CTGCATTGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATCCAACTTGAA 124
      |||||||
QY 712 CATTCTACTTTTCTTGATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGA 771
      |||||||
Db 123 CATTCTACTTTTCTTGATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGA 64
      |||||||
QY 772 GTTGTCTAGCTTTTAACTTTATATCAATCTTTTTCAGATTAACATAATTTGTGATG 831
      |||||||
Db 63 GTTGTCTAGCTTTTAACTTTATATCAATCTTTTTCAGATTAACATAATTTGTGATG 4
      |||||||
QY 832 GTG 834
      |||
Db 3 GTG 1

RESULT 7
AI283262/c
LOCUS
DEFINITION
  AI283262
  qk50g08.x1 NCI-CGAP_Co8 Homo sapiens cDNA clone IMAGE:1872446 3'
  similar to WP:F26A3.4 CE09669 PROTEIN-TYROSINE PHOSPHATASE ;, mRNA
  sequence.
ACCESSION
  AI283262
VERSION
  AI283262.1 GI:3921495
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 441)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www.bio.llnl.gov/bbrp/image/image.html
  Insert Length: 1361 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 404.
  Location/Qualifiers
    1..441
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1872446"
    /clone_lib="NCI-CGAP_Co8"
    /tissue_type="adenocarcinoma"
    /lab_host="DH10B"
    /note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
    modified polylinker; 1st strand cDNA was prepared from
    colon adenocarcinoma, and was then primed with a Not I -
    oligo(dT) primer. Double-stranded cDNA was ligated to Eco
    RI adaptors (Pharmacia), digested with Not I and cloned
    into the Not I and Eco RI sites of the modified pT73
    vector. Library is normalized. Library was constructed by
    Bento Soares and M. Fatima Bonaldo."
  109 a 103 c 122 g 107 t

BASE COUNT
ORIGIN

```



```

FEATURES
source
Location/Qualifiers
1..446
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="NT0290"
    /dev_stage="Adult"
    /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
    Site_2: SmaI; A mini-library was made by cloning products
    derived from ORESTES PCR (U.S. Letters Patent application
    No. 196,716 - Ludwig Institute for Cancer Research)
    profiles into the pUC 18 vector. Reverse transcription of
    tissue mRNA and cDNA amplification were performed under
    low stringency conditions."
97 a 141 c 107 g 101 t
BASE COUNT
        97
ORIGIN

```

[illegible]

```

Qy 518 GCTCATCCA 526
|||||
Db 438 GCTCATCCA 446

RESULT 10
BQ934977
LOCUS BQ934977 894 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8981260 NCI_CGAP_Co24 Mus musculus cDNA clone
IMAGE:6396450 5', mRNA sequence.
BQ934977
ACCESSION BQ934977.1 GI:22350360
VERSION BQ934977
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 894)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13892 row: 1 column: 19
High quality sequence stop: 676.
FEATURES
source
1..894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6396450"
/clone_lib="NCI_CGAP_Co24"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 229 a 229 c 220 g 216 t
ORIGIN
Query Match 43.1%; Score 359.8; DB 14; Length 894;
Best Local Similarity 81.9%; Pred. No. 8e-89;
Matches 415; Conservative .0; Mismatches 92; Indels 0; Gaps 0;

Qy 152 TATCAGCAATGGTGTGCGCCGACACACACAGCTCATGCTGTCTAGGACACAGATCACCAT 211
|||||
Db 2 TATCAGCAATGGTGTGCTGCCAACACACAGCTCTCTACTGTCTCCAGCAATCAGATCACCAC 61

Qy 212 GGTATCAATGCTCTCAGTGGAGTAGTGAACACCTCTGTATGAGATATCCAGTACATGCA 271
|||||
Db 62 AGTCATCAACGCTCTCAGTAGAGGTAGCAACACCTCTTACGAGATATCCAGTATGTGCA 121

Qy 272 GGTACCTGTGGCTGACTCCCTCACTACCTACCTGCTCTGTGACTTCTTTGACCCCTATTGCTGA 331
|||||
Db 122 GGTGCTGTGTGCTGCTGCGCCGTCGACCGCTCTCTCAATTTCTTCGATTCGGTTCGCGA 181

Qy 332 CCATATCCACAGCTGTGAGATGAGCAGCGCTACTTCTGCTGACATGCTGCTGTGTGT 391
|||||
Db 182 CCGTATCCATTCGGTGGAGATGAGAGAGGCGCCGACACTGTTCGATTTGCTGCTGCGGGT 241

Qy 392 GAGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
|||||
Db 242 GAGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

Qy 452 GGACGCGCCACAGCTGGACCAAGTATCATGCCGCGCCCATCATCCGACCCCAACAGCGGCTTTTG 511
|||||

```

```

Db 302 AGATGCCACACTTGGACCAAGTCATCGAGCCCATCATCGGCCCAACAGTGGCTTTTG 361
|||||
Qy 512 GGACGAGCTCATCCACTATGATTCCTCAATGTTTGGGAAGACACTGTGCACATGGTCAG 571
|||||
Db 362 GGACGAGCTCATCCATTCAGGATTCGAGCTCTTTTGGCAAGAAATACAATGCAGATGATGA 421
|||||
Qy 572 TTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGAAAGTCCGTTTGATGATTCCACT 631
|||||
Db 422 CTCGCCAATGGGAGGATCCAGACATCTAGGAGAAGAGACCCGTTTGATGATCCCACT 481
|||||
Qy 632 GTGAGCCATCCACGAGCCCTCCCTGCAAT 658
|||||
Db 482 GTAAGCCAGCCACACAGTAACCTCCAGT 508

RESULT 11
BQ333140
LOCUS BQ333140 679 bp mRNA linear EST 27-FEB-2001
DEFINITION 602430837F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4548585 5',
mRNA sequence.
ACCESSION BQ333140
VERSION BQ333140.1 GI:13139578
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 679)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CMI238 row: 1 column: 10
High quality sequence stop: 679.
FEATURES
source
1..679
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4548585"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOT87; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 145 a 195 c 187 g 151 t 1 others
ORIGIN
Query Match 42.7%; Score 356.4; DB 12; Length 679;
Best Local Similarity 96.8%; Pred. No. 6.8e-88;
Matches 363; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 23 CATTTTGGTGTTCGCTGACTGCTGACCACTGACCCACCCTTGTGATGACAGCACCTC 82
|||||
Db 276 CAGCCTTGGTGTTCGCTGACTGCTGACCACTGACCCACCNGCTTGTGATGACAGCACCTC 335
|||||
Qy 83 GTGTGCTTCCCACTTCCGTCAGTTCGGCAGCCCTCAGTCAGCGGCTCTCGCAGATAACCA 142
|||||
Db 336 GTGTGCTTCCCACTTCCGTCAGTTCGGCAGCCCTCAGTCAGCGGCTCTCGCAGATAACCA 395
|||||

```

[illegible]

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RESULT     | 313                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| BI907369   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| LOCUS      | 744 bp mRNA linear EST 16-OCT-2001                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| DEFINITION | GI03063702F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212753 5',                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| ACCESSION  | BI907369                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| VERSION    | BI907369.1 GI:16170196                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| KEYWORDS   | mRNA sequence.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| SOURCE     | EST.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| ORGANISM   | human.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| REFERENCE  | Homo sapiens                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| AUTHORS    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| TITLE      | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| JOURNAL    | 1 (bases 1 to 744)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| COMMENT    | NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .<br>National Institutes of Health, Mammalian Gene Collection (MGC)<br>Unpublished (1999)<br>Contact: Robert Strausberg, Ph.D.<br>Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a><br>Tissue Procurement: Life Technologies, Inc.<br>cDNA Library Preparation: Life Technologies, Inc.<br>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)<br>DNA Sequencing by: Incyte Genomics, Inc.<br>Clone distribution: MGC clone distribution information can be<br>found through the I.M.A.G.E. Consortium/LLNL at:<br><a href="http://image.llnl.gov">http://image.llnl.gov</a><br>Plate: L1AM1534 row: d column: 02<br>High quality sequence stop: 743. |

```

FEATURES
  source
    Location/Qualifiers
      1. 744
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5212753"
        /clone_lib="NIH_MGC_l18"
        /tissue_type="leukocyte"
        /lab_host="DH10B"
        /note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV
        (destroyed); RNA source leukocytes from anonymous pool of
        non-activated adult donors. Library is oligo-dT primed
        and directionally cloned (EcoRV site is destroyed upon
        cloning). Average insert size 1.7 kb, insert size range
        1.2-3.3 kb. Library is normalized and enriched for
        full-length clones and was constructed by C. Gruber
        (Invitrogen). Research Genetics tracking code 027. Note:
        this is a NIH_MGC Library."
        152 a      222 c      216 g      154 t
BASE COUNT
ORIGIN
  Query Match      38.9%; Score 324.6; DB 13; Length 744;
  Best Local Similarity 98.8%; Pred. No. 4.6e-79;
  Matches 327; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QV 23 CATTCTTGGTGTTCGGCTGACTGCTGACCACCTGACCACCGCCCTTGATGACGACCCCTC 82

```

Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamly, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

5 (bases 1 to 840)  
A4chi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saigo, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejlina, Y., Toyota, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCCAAGGACGCTTTTTTTTTTTTTTTTNN 3'], cDNA was prepared by using tris(hydroxymethyl)aminomethane reverse transcriptase and subsequently prepared for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGCGCCCAATTATCTCGAGTTAATTAATTCCTCCCGCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

end. Location/Qualifiers  
1. .840  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:1700094E07"  
/db\_xref="MG:1906548"  
/db\_xref="taxon:10090"  
/clone="1700094E07"  
/sex="male"  
/tissue\_type="testis"  
/cdone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
93. .662  
/note="data source:SPTR, source key:Q9H596, evidence:ISS homolog to BA386N14.1 (NOVEL PROTEIN SIMILAR TO A DUAL SPECIFICITY PHOSPHATASE)  
putative"  
/codon\_start=1  
/protein\_id="BAB24847.1"

CDS



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:45:03 ; Search time 1834 Seconds  
(without alignments)  
13234.318 Million cell updates/sec

Title: US-09-527-376-1  
Perfect score: 834  
Sequence: 1 ctttctgtatttttgc.....aaactaatgtgagatggtg 834

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID        | Description |
|------------|-------|-------|--------|----|-----------|-------------|
| 1          | 805.6 | 96.6  | 1407   | 6  | AX202236  | Sequence    |
| 2          | 805.6 | 96.6  | 2454   | 9  | AK056074  | Homo sapi   |
| 3          | 804   | 96.4  | 39401  | 9  | AC003072  | Human BAC   |
| 4          | 704.6 | 84.5  | 775    | 6  | AX086024  | Sequence    |
| 5          | 647   | 77.6  | 1865   | 9  | BC004110  | Homo sapi   |
| 6          | 640   | 76.7  | 1335   | 9  | BC030987  | Homo sapi   |
| 7          | 575.6 | 69.0  | 2921   | 9  | BC028724  | Homo sapi   |
| 8          | 474.4 | 56.9  | 191511 | 2  | AC125764  | Rattus no   |
| 9          | 445   | 53.4  | 1655   | 9  | AK025036  | Homo sapi   |
| 10         | 423.6 | 50.8  | 2681   | 10 | BC020036  | Mus muscu   |
| 11         | 423.6 | 50.8  | 186768 | 2  | AL807241  | Mus muscu   |
| 12         | 423.6 | 50.8  | 205509 | 10 | AL731853  | Mouse DNA   |
| 13         | 369.8 | 44.3  | 908    | 6  | AX086022  | Sequence    |
| 14         | 362.6 | 43.5  | 140717 | 9  | AL133545  | Human DNA   |
| 15         | 362.4 | 43.5  | 935    | 6  | AX040320  | Sequence    |
| 16         | 343.4 | 41.2  | 889    | 9  | AF143321  | Homo sapi   |
| 17         | 343.4 | 41.2  | 889    | 9  | AY121807  | Homo sapi   |
| 18         | 322.6 | 38.7  | 177327 | 2  | AC113681  | Rattus no   |
| 19         | 322.6 | 38.7  | 189620 | 2  | AC126908  | Rattus no   |
| 20         | 313.4 | 37.6  | 161816 | 10 | AL773547  | Mouse DNA   |
| 21         | 187.2 | 22.4  | 1425   | 10 | BC002130  | Mus muscu   |
| 22         | 187.2 | 22.4  | 196812 | 10 | AL645615  | Mouse DNA   |
| 23         | 187.2 | 22.4  | 201430 | 2  | AL596447  | Mus muscu   |
| 24         | 185.6 | 22.3  | 597    | 10 | AF120113  | Mus muscu   |
| 25         | 177.6 | 21.3  | 1167   | 9  | AF120032  | Homo sapi   |
| 26         | 177.6 | 21.3  | 1445   | 9  | BC004448  | Homo sapi   |
| 27         | 177.6 | 21.3  | 1454   | 9  | BC000370  | Homo sapi   |
| 28         | 177.6 | 21.3  | 1471   | 9  | AF038844  | Homo sapi   |
| 29         | 177.6 | 21.3  | 1474   | 9  | BC001894  | Homo sapi   |
| 30         | 177.6 | 21.3  | 1729   | 6  | AR122778  | Sequence    |
| 31         | 177.6 | 21.3  | 136222 | 9  | AC004099  | Homo sapi   |
| 32         | 151   | 18.1  | 461    | 11 | G28106    | human STS   |
| 33         | 137.6 | 16.5  | 2259   | 9  | AK027210  | Homo sapi   |
| 34         | 98.2  | 11.8  | 142093 | 9  | AC026884  | Homo sapi   |
| 35         | 98.2  | 11.8  | 170934 | 2  | AL354684  | Homo sapi   |
| 36         | 90    | 10.8  | 1059   | 5  | AF026522  | Gallus ga   |
| 37         | 88.8  | 10.6  | 1933   | 6  | AX306137  | Sequence    |
| 38         | 88.8  | 10.6  | 1933   | 10 | MM3CH134M | Mouse mRNA  |
| 39         | 88.8  | 10.6  | 1941   | 10 | BC006967  | Mus muscu   |
| 40         | 88.2  | 10.6  | 1467   | 5  | AF167296  | Gallus ga   |
| 41         | 88    | 10.6  | 2000   | 6  | AX337238  | Sequence    |
| 42         | 88    | 10.6  | 2000   | 6  | AX411130  | Sequence    |
| 43         | 88    | 10.6  | 2000   | 9  | HSCL100   | H.sapiens C |
| 44         | 88    | 10.6  | 2026   | 9  | BC022463  | Homo sapi   |
| 45         | 87    | 10.4  | 1209   | 6  | AX369208  | Sequence    |

ALIGNMENTS

RESULT 1  
AX202236  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AX202236  
Sequence 15 from Patent WO0153469.  
AX202236  
AX202236.1 GI:15392011  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1407)  
Bandman,O., Tang,Y.T., Azimzai,Y., Yue,H., Baughn,M.R.,  
Hillman,J.L., Lal,P., Wang,E., Gandhi,A.R., Policky,J.L. and  
Mathur,P.

AX202236  
Sequence 15 from Patent WO0153469.  
AX202236  
AX202236.1 GI:15392011  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1407)  
Bandman,O., Tang,Y.T., Azimzai,Y., Yue,H., Baughn,M.R.,  
Hillman,J.L., Lal,P., Wang,E., Gandhi,A.R., Policky,J.L. and  
Mathur,P.

AX202236  
Sequence 15 from Patent WO0153469.  
AX202236  
AX202236.1 GI:15392011  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1407)  
Bandman,O., Tang,Y.T., Azimzai,Y., Yue,H., Baughn,M.R.,  
Hillman,J.L., Lal,P., Wang,E., Gandhi,A.R., Policky,J.L. and  
Mathur,P.

## RESULT 2

```
QY 263 GTACATGAGGTTACCTGTGGTGAATCCCTAACTCACTCTCTGTGACTCTTTTGACCC 322
Db 696 GTACATGAGGTTACCTGTGGTGAATCCCTAACTCACTCTCTGTGACTCTTTTGACCC 755
QY 323 TATTGCTGACCATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGCTGCACTGTC 382
Db 756 TATTGCTGACCATCCACAGCGTGGAGATGAAGCAGGCGGTACTTTGCTGCACTGTC 815
QY 383 TGCTGGTGTGAGCCGCTCAGCTGCCCTGTGGCTCCCTACCTCATGAAGTACCAAGCCCAT 442
Db 816 TGCTGGTGTGAGCCGCTCAGCTGCCCTGTGGCTCCCTACCTCATGAAGTACCAAGCCCAT 875
QY 443 GTCCCTGCTGGAGCCGCTCAGCTGCCCTGTGGCTCCCTACCTCATGAAGTACCAAGCCCAT 502
Db 876 GTCCCTGCTGGAGCCGCTCAGCTGCCCTGTGGCTCCCTACCTCATGAAGTACCAAGCCCAT 935
QY 503 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCGAATTTGGCAAGACACTGTGCA 562
Db 936 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCGAATTTGGCAAGACACTGTGCA 995
QY 563 CATGCTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 622
Db 996 CATGCTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 1055
QY 623 GATTCACCTGTGAGCCATCCACGAGCCCTGCAATGAGTGCAGAGGTACAGATCTATTG 682
Db 1056 GATTCACCTGTGAGCCATCCACGAGCCCTGCAATGAGTGCAGAGGTACAGATCTATTG 1115
QY 683 TTGATCTTACACCAAGATCCAACTTGAACATCTTACCTTTGTTGATACAGAAAAACA 742
Db 1116 TTGATCTTACACCAAGATCCAACTTGAACATCTTACCTTTGTTGATACAGAAAAACA 1175
QY 743 GATGATGCTTTTATGACACAAAAAGATTGCTGTAGCTTTTAACTTTTATATCCATT 802
Db 1176 GATGATGCTTTTATGACACAAAAAGATTGCTGTAGCTTTTAACTTTTATATCCATT 1235
QY 803 TTTTTCAGATTAATACTAATTGTGAGATGGTG 834
Db 1236 TTTTTCAGATTAATACTAATTGTGAGATGGTG 1267

RESULT 3
AC003072/c 39401 bp DNA linear PRI 04-FEB-2000
LOCUS Human BAC clone CTA-963H5 from 22q12.1-qter, complete sequence.
DEFINITION AC003072
ACCESSION AC003072
VERSION AC003072.1 GI:2588642
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 39401)
AUTHORS Murray,J., Wohlmann,P., Bauer,C. and Courtney,L.
TITLE The sequence of H. sapiens BAC clone CTA-963H5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 39401)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1997) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 3 (bases 1 to 39401)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 4 (bases 1 to 39401)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
```

Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
----- Summary Statistics  
Center project name: H\_BR963H05  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre chromosome 22 mapping group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22/>

#### SOURCE INFORMATION:

This clone is from the human BAC library described by U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).  
VECTOR: pBelBAC11  
Selection: chloramphenicol

#### NEIGHBORING SEQUENCE INFORMATION:

The clone being sequenced to the left is RP1-56J10. The actual start of this clone is at base position 1 of CTA-963H5. This clone is part of an unanchored island, orientation is unknown.

| FEATURES      | Source                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| repeat_region | 1. 39401<br>/organism="Homo sapiens"<br>/db_xref="taxon:9606"<br>/chromosome="22"<br>/map="22q12.1-qter"<br>/clone="CTA-963H5"<br>/clone_lib="BKB"<br>complement(1142..1435)<br>/rpt_family="ALU"<br>complement(1750..2045)<br>/rpt_family="ALU"<br>2937..3203<br>/rpt_family="ALU"<br>complement(3758..4038)<br>/rpt_family="ALU"<br>4595..4766<br>/rpt_family="ALU"<br>4915..5230<br>/rpt_family="ALU"<br>5239..5363<br>/rpt_family="ALU"<br>5385..5675<br>/rpt_family="ALU"<br>5692..5875<br>/rpt_family="ALU"<br>6019..6148<br>/rpt_family="ALU"<br>6625..6895<br>/rpt_family="ALU"<br>complement(7143..7413)<br>/rpt_family="ALU"<br>complement(8074..8235)<br>/rpt_family="ALU"<br>complement(8610..8902)<br>/rpt_family="ALU"<br>complement(9218..9360) |





```

|||||
Db 310 CATCCGACCAACAGCGGCTTTGGGAGCAGCTCATCCACTATGAGTCCAAATTTGTTGG 369
QY 548 CAAGAACACATGTCACATGTCAGTCCCTCCAGTGGGAATCATCCCTGACATCATGAGAA 607
Db 370 CAAGAACACATGTCACATGTCAGTCCCTCCAGTGGGAATCATCCCTGACATCATGAGAA 429
QY 608 GGAAGTCGCTTTGATGATTCACATGTCAGGATCCACACGAGCCCTGCAATGGAGTCAGA 667
Db 430 GGAAGTCGCTTTGATGATTCACATGTCAGGATCCACACGAGCCCTGCAATGGAGTCAGA 489
QY 668 GGTACAGATCTATGTTGATCTTACACCAAGATCCAAACTTGAACTTACATTTTGTGTTG 727
Db 490 GGTACAGATCTATGTTGATCTTACACCAAGATCCAAACTTGAACTTACATTTTGTGTTG 549
QY 728 ATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGAGTGTGCTAGCTTTTA 787
Db 550 ATACAGAAAAACAGATGATGCTTTTATGAGCACAAAAAGAGTGTGCTAGCTTTTA 609
QY 788 ACTTTATAATCCATTTTTTTTCAGATTAATACTAATGTGAGATGGTG 834
Db 610 ACTTTATAATCCATTTTTTTTCAGATTAATACTAATGTGAGATGGTG 656

```

```

RESULT 6
BC030987
LOCUS
DEFINITION
Homo sapiens, similar to RIKEN CDNA 4930527G07 gene, clone
MGC:32658 IMAGE:4730282, mRNA, complete cds.
ACCESSION
BC030987
VERSION
BC030987.1 GI:21410213
KEYWORDS
MGC.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1335)
Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxill.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 41 Row: m Column: 12

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction, Similarity but not identity to protein.

```

FEATURES
source
1. .1335
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:32658 IMAGE:4730282"
/tissue_type="placenta"
/clone_lib="NIH_MGC_79"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
460. .1026
CDS

```

```

/codon_start=1
/product="Similar to RIKEN CDNA 4930527G07 gene"
/protein_id="AAH30987.1"
/db_xref="GI:21410214"
/translation="MTAPSCAFVQFQPSVSGLSQITKSLYISNGVAANKMLLSSN
QITMWINSEVNTLIEDIQMVPVADSPNSRUCDFEDPIADHLSHSEVMQGRTELL
HCAAGVSRSAALCLAYLMKYHAMSLLDAHTWTKSRPIIRPNSGFWEQLIHFEQLFG
KNTVMVSSPVGMIPDIYEKEVRLMPL"
BASE COUNT 304 a 399 c 346 g 286 t
ORIGIN

Query Match 76.7%; Score 640; DB 9; Length 1335;
Best Local Similarity 98.5%; Pred. No. 2.6e-164;
Matches 646; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 23 CATTCTTGGTGTTCCTGACTGCTGACCACTGACCCACCGCTTGTGATGACAGCACCCCTC 82
Db 414 CAGCCTTCGTGCTTCGCTGACTGCTGACCACTGACCCACCGCTTGTGATGACAGCACCCCTC 473
QY 83 GTGTGCTTCCCACTGCTGAGTTCAGTTCGCGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAA 142
Db 474 GTGTGCTTCCCACTGCTGAGTTCAGTTCGCGGAGCCCTCAGTCAGCGGCTCTCGCAGATAACCAA 533
QY 143 AAGCCTGTATATCAGCAATGGTGTGGCCGCCAACACAAAGCTCATGCTGTCTAGCAACCA 202
Db 534 AAGCCTGTATATCAGCAATGGTGTGGCCGCCAACACAAAGCTCATGCTGTCTAGCAACCA 593
QY 203 GATCACCATTGGTTCATCAATGCTCTCAGTGGAGGTAGTGAACACCTTGTATGAGATATCCA 262
Db 594 GATCACCATTGGTTCATCAATGCTCTCAGTGGAGGTAGTGAACACCTTGTATGAGATATCCA 653
QY 263 GTACATGCAGGTACCTGTGGCTGACTGCTCCCTAACTACGCTCTCTGTGACATCTCTTTGACCC 322
Db 654 GTACATGCAGGTACCTGTGGCTGACTGCTCCCTAACTACGCTCTCTGTGACATCTCTTTGACCC 713
QY 323 TATTGTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCGTACTTGTGTCAGCTGTGC 382
Db 714 TATTGTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCGTACTTGTGTCAGCTGTGC 773
QY 383 TGTGCTGTGAGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
Db 774 TGTGCTGTGAGCGGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
QY 443 GTCCTCTGCTGGAGCGCCACACACGCTGGAGTCAATGCGGCGCCCATCATCGAGCCCAACAG 502
Db 834 GTCCTCTGCTGGAGCGCCACACACGCTGGAGTCAATGCGGCGCCCATCATCGAGCCCAACAG 893
QY 503 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCTCAATTTTGGCAAGAACACTGTGCA 562
Db 894 CGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCTCAATTTTGGCAAGAACACTGTGCA 953
QY 563 CATGGTCAGTTCCTCCAGTGGGAATGATCCCTGACATCTATGAGAGGAAGTCCGTTTGTAT 622
Db 954 CATGGTCAGTTCCTCCAGTGGGAATGATCCCTGACATCTATGAGAGGAAGTCCGTTTGTAT 1013
QY 623 GATTCCACTGTGAGCGCATCCACGAGCCCTGCAATGGAGTTCAGAGTACAGATCT 678
Db 1014 GATTCCACTGTGAGCGCATCCACGAGCCCTGCAATGGAGTTCAGAGTTCAGAGTTCGCTT 1069

RESULT 7
BC028724
LOCUS
DEFINITION
Homo sapiens, similar to Unknown (protein for IMAGE:3689593), clone
MGC:33231 IMAGE:5269586, mRNA, complete cds.
ACCESSION
BC028724
VERSION
BC028724.1 GI:20381112
KEYWORDS
MGC.
SOURCE
Homo sapiens.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2921)
REFERENCE
1 (bases 1 to 2921)

```





```

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwono,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pichens,R., Primus,E., Fu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 191511)
Worley,K.C.
Direct Submission
Submitted (30-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 191511)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 2, 2002 this sequence version replaced gi:21630358.
-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GZFB
Center clone name: CH230-229p3
-----
Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 143602 bases at least Q40
Consensus quality: 151894 bases at least Q30
Consensus quality: 158330 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1316: contig of 1316 bp in length
* 1317: 1416: gap of unknown length
* 1417: 3055: contig of 1639 bp in length
* 3056: 3155: gap of unknown length
* 3156: 4162: contig of 1007 bp in length
* 4163: 4262: gap of unknown length
* 4263: 5639: contig of 1377 bp in length
* 5640: 5739: gap of unknown length
* 5740: 6782: contig of 1043 bp in length
* 6783: 6882: gap of unknown length
* 6883: 8022: contig of 1140 bp in length
* 8023: 8122: gap of unknown length
* 8123: 9761: contig of 1639 bp in length
* 9762: 9861: gap of unknown length
* 9862: 11958: contig of 2097 bp in length
* 11959: 12038: gap of unknown length
*
* 12059: 13754: contig of 1696 bp in length
* 13755: 13854: gap of unknown length
* 13855: 15390: contig of 1536 bp in length
* 15391: 15490: gap of unknown length
* 15491: 16642: contig of 1152 bp in length
* 16643: 16742: gap of unknown length
* 16743: 18412: contig of 1670 bp in length
* 18413: 18512: gap of unknown length
* 18513: 20412: contig of 1900 bp in length
* 20413: 20512: gap of unknown length
* 20513: 22041: contig of 1529 bp in length
* 22042: 22141: gap of unknown length
* 22142: 23649: contig of 1508 bp in length
* 23650: 23749: gap of unknown length
* 23750: 25360: contig of 1611 bp in length
* 25361: 25460: gap of unknown length
* 25461: 27686: contig of 2226 bp in length
* 27687: 27786: gap of unknown length
* 27787: 29677: contig of 1891 bp in length
* 29678: 29777: gap of unknown length
* 29778: 32105: contig of 2328 bp in length
* 32106: 34956: contig of 2751 bp in length
* 34957: 35056: gap of unknown length
* 35057: 37267: contig of 2211 bp in length
* 37268: 37367: gap of unknown length
* 37368: 39228: contig of 1861 bp in length
* 39229: 39328: gap of unknown length
* 39329: 40783: contig of 1455 bp in length
* 40784: 40883: gap of unknown length
* 40884: 42941: contig of 2058 bp in length
* 42942: 43041: gap of unknown length
* 43042: 44879: contig of 1838 bp in length
* 44880: 44879: gap of unknown length
* 44880: 46226: contig of 1247 bp in length
* 46227: 46326: gap of unknown length
* 46327: 48011: contig of 1685 bp in length
* 48012: 48111: gap of unknown length
* 48112: 49971: contig of 1860 bp in length
* 49972: 50071: gap of unknown length
* 50072: 51849: contig of 1778 bp in length
* 51850: 51949: gap of unknown length
* 51950: 54291: contig of 2342 bp in length
* 54292: 54391: gap of unknown length
* 54392: 57401: contig of 3010 bp in length
* 57402: 57501: gap of unknown length
* 57502: 59171: contig of 1670 bp in length
* 59172: 59271: gap of unknown length
* 59272: 62524: contig of 3253 bp in length
* 62525: 62624: gap of unknown length
* 62625: 64053: contig of 1429 bp in length
* 64054: 64153: gap of unknown length
* 64154: 66493: contig of 2340 bp in length
* 66494: 66593: gap of unknown length
* 66594: 68444: contig of 1851 bp in length
* 68445: 68544: gap of unknown length
* 68545: 71479: contig of 2935 bp in length
* 71480: 71579: gap of unknown length
* 71580: 74791: contig of 3212 bp in length
* 74792: 74891: gap of unknown length
* 74892: 78202: contig of 3311 bp in length
* 78203: 78302: gap of unknown length
* 78303: 82035: contig of 3733 bp in length
* 82036: 82135: gap of unknown length
* 82136: 85344: contig of 3209 bp in length
* 85345: 85444: gap of unknown length
* 85445: 88700: contig of 3256 bp in length
* 88701: 88800: gap of unknown length
* 88801: 93392: contig of 4492 bp in length
* 93393: 93392: gap of unknown length
* 93393: 98296: contig of 4904 bp in length
* 98297: 98396: gap of unknown length
* 101597: contig of 3201 bp in length

```







Consensus quality: 186504 bases at least Q20  
 Insert size: 186668; sum-of-contigs  
 Insert size: 195161; 5.7% error; agarose-fp  
 Quality coverage: 8.45x in Q20 bases; sum-of-contigs Quality  
 coverage: 8.17x in Q20 bases; agarose-fp

\*\*\*\*\*  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 143379: contig of 143379 bp in length  
 \* 143380 143479: gap of 100 bp  
 \* 143480 186768: contig of 43289 bp in length.

#### FEATURES

source  
 Location/Qualifiers  
 1. 186768  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone\_lib="RPCI-23"  
 /clone="RP23-453E10"  
 1. 143379  
 /note="assembly\_fragment:02056.0"  
 143480..186768  
 /note="assembly\_fragment:02056.1"  
 /note="assembly\_fragment:02056.1"  
 BASE COUNT 50057 a 44175 c 44369 g 48067 t 100 others  
 ORIGIN

Query Match 50.8%; Score 423.6; DB 2; Length 186768;  
 Best Local Similarity 81.8%; Pred. No. 1.5e-104;  
 Matches 489; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 61 CCGCTTTGATGACAGCACCCTCGTGTGCTTCCAGTTTCCGCGGAGCCCTCAGTCA 120  
 Db 88221 CCACCTTTGATGACATCACCGTGGAGTGCCTTCCAGTTTCCGCGGAGCCCTCAGTCA 88280

Qy 121 CGGCGCTCTCCAGATACCAAAAGCCGTGTATATCAGCAATGGTGTGGCCGCCAACACA 180  
 Db 88281 GAGCGCTCTCCAGATACCAAAAGCCGTGTATATCAGCAATGGTGTGGCCGCCAACACA 88340

Qy 181 AGCTCATGCTCTAGCAACAGATCACATGGTGCATCAATGTCTCAGTGGAGTAGTGA 240  
 Db 88341 AGCTCATGCTCTAGCAACAGATCACATGGTGCATCAATGTCTCAGTGGAGTAGTGA 88400

Qy 241 ACACCTTTGATGAGATATCCAGTACATGCAGTACCTGTGGCTGACTCCCTCAACTCAC 300  
 Db 88401 ACACCTTTGATGAGATATCCAGTACATGCAGTACCTGTGGCTGACTCCCTCAACTCAC 88460

Qy 301 GTCCTGTGACTCTTTGACCCCTATTTGTGACATATCCACAGCGTGGAGATGAAGCAG 360  
 Db 88461 GTCCTGTGACTCTTTGACCCCTATTTGTGACATATCCACAGCGTGGAGATGAAGCAG 88520

Qy 361 GCGGTACTTTGTCGACTGCTGCTGTGTGGAGCGCTGAGTGGCTGTGGCTGCGCT 420  
 Db 88521 GCGGTACTTTGTCGACTGCTGCTGTGTGGAGCGCTGAGTGGCTGTGGCTGCGCT 88580

Qy 421 ACCTCATGAAGTACACGCCATGTCCTGTGGAGCGGCCACACCTGGACCAAGTCATGCC 480  
 Db 88581 ACCTCATGAAGTACACGCCATGTCCTGTGGAGCGGCCACACCTGGACCAAGTCATGCC 88640

Qy 481 GGCCATCATCCGACCAACAGCGCTTTTGGAGCAGCTCATCCACTATGAGTTCCAA 540  
 Db 88641 GGCCATCATCCGACCAACAGCGCTTTTGGAGCAGCTCATCCACTATGAGTTCCAA 88700

Qy 541 TGTGTGCAAGAACTGTGCACATGGTCACTTCCCACTGGAGTATCCCTGACATCT 600  
 Db 88701 TGTGTGCAAGAACTGTGCACATGGTCACTTCCCACTGGAGTATCCCTGACATCT 88760

Qy 601 ATGAGAAGGAAGTCGGTTTGTATGATTCACATGTGAGGCATCCCAAGCCCTCGATT 658  
 Db 88761 ATGAGAAGGAAGTCGGTTTGTATGATTCACATGTGAGGCATCCCAAGCCCTCGATT 88820

Db 88761 ACGAGAAGGAGACCGTTTGTATGATCCACTGTAAAGCCAGCCACCACTAATCCAGT 88818

#### RESULT 12

AL731853

LOCUS

DEFINITION

Mouse DNA sequence from clone RP23-12865 on chromosome 11, complete

sequence.

AL731853

ACCESSION

AL731853.9

GI:21953059

HTG.

KEYWORDS

SOURCE

ORGANISM

Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 205509)

Dunn.M.

Direct Submission

Submitted (19-JUL-2002)

Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk

On Jul 24, 2002 this sequence version replaced gi:21540128.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

-----

COMMENT

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em, EMBL; Sw,

SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep

RP23-12865 is

constructed by the group of Pieter de Jong.

For further details see http://www.chori.org/bacpac/home.htm

VECTOR: pBACE3.6.

Location/Qualifiers

1. 205509

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="11"

/clone="RP23-12865"

/clone\_lib="RPCI-23"

BASE COUNT 55286 a 49833 c 49670 g 50720 t

ORIGIN

Query Match 50.8%; Score 423.6; DB 10; Length 205509;

Best Local Similarity 81.8%; Pred. No. 1.5e-104;

Matches 489; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 61 CCGCTTTGATGACAGCACCCTCGTGTGCTTCCAGTTTCCGCGGAGCCCTCAGTCA 120

Db 188011 CCACCTTTGATGACATCACCGTGGAGTGCCTTCCAGTTTCCGCGGAGCCCTCAGTCA 188070

Qy 121 CGGCGCTCTCCAGATACCAAAAGCCGTGTATATCAGCAATGGTGTGGCCGCCAACACA 180

Db 188071 GAGCGCTCTCCAGATACCAAAAGCCGTGTATATCAGCAATGGTGTGGCCGCCAACACA 188130

Qy 181 AGCTCATGCTCTAGCAACAGATCACATGGTGCATCAATGTCTCAGTGGAGTAGTGA 240



This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>

RP11-386N14 is from the library RPCI-11.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-386N14 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true right end of clone RP11-386N14 is at 140717 in this sequence. The true right end of clone RP6-105D16 is at 100 in this sequence.

## FEATURES

## source

```
1. 140717
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="p11.23-11.4"
/clone="RP11-386N14"
/clone_lib="RPCI-11.2"
```

```
repeat_region
1. 200
/feature="AluY repeat: matches 101. .293 of consensus"
326. .633
/feature="AluY repeat: matches 5. .310 of consensus"
717. .941
/feature="MER20 repeat: matches 3. .217 of consensus"
complement(1062. 1294)
/feature="match: GSS: Em:AQ629580"
1062. 1256
/feature="AluSg/x repeat: matches 117. .310 of consensus"
complement(1062. 1245)
/feature="match: GSS: Em:AQ215148"
complement(1062. 1243)
/feature="match: GSS: Em:AQ005454"
1066. 1245
/feature="match: GSS: Em:AQ487216"
1294. 1404
/feature="MER3 repeat: matches 50. .165 of consensus"
1429. 1580
/feature="AluY repeat: matches 134. .285 of consensus"
1581. 1884
/feature="AluY repeat: matches 2. .303 of consensus"
1887. 1925
/feature="MER3 repeat: matches 159. .205 of consensus"
1900. 1940
/feature="MER3 repeat: matches 37. .77 of consensus"
1942. 2090
/feature="MIR repeat: matches 29. .172 of consensus"
2091. 2377
/feature="AluSx repeat: matches 1. .294 of consensus"
2378. 2437
/feature="MIR repeat: matches 172. .232 of consensus"
2644. 2958
/feature="AluSp repeat: matches 1. .313 of consensus"
3022. 3323
/feature="AluSx repeat: matches 1. .300 of consensus"
3496. 3581
/feature="L1M5 repeat: matches 7932. .7925 of consensus"
3698. 3743
/feature="23 copies 2 mer tt 73% conserved"
```

```
repeat_region
3749. .3969
/feature="AluJo repeat: matches 82. .296 of consensus"
4090. 4207
/feature="L1P repeat: matches 3626. .3747 of consensus"
4228. 4273
/feature="23 copies 2 mer tt 73% conserved"
4276. 4625
/feature="L1PA5 repeat: matches 5795. .6145 of consensus"
4626. 4687
/feature="L1PA5 repeat: matches 5744. .5805 of consensus"
4802. 5089
/feature="AluSp repeat: matches 1. .289 of consensus"
5528. 5820
/feature="AluSg repeat: matches 1. .293 of consensus"
6003. 6306
/feature="AluSg repeat: matches 1. .304 of consensus"
6325. 6618
/feature="AluY repeat: matches 1. .298 of consensus"
6890. 7198
/feature="AluSx repeat: matches 1. .308 of consensus"
7303. 7611
/feature="AluSg repeat: matches 1. .309 of consensus"
7669. 7965
/feature="AluSx repeat: matches 1. .305 of consensus"
7984. 8110
/feature="FLAM_C repeat: matches 1. .127 of consensus"
8237. 8405
/feature="AluSx repeat: matches 124. .294 of consensus"
8406. 8714
/feature="AluY repeat: matches 1. .308 of consensus"
9006. 9517
/feature="L1MA5A repeat: matches 5773. .6291 of consensus"
10250. 10558
/feature="AluSg repeat: matches 3. .305 of consensus"
10560. 10695
/feature="AluSx repeat: matches 1. .136 of consensus"
complement(10653. 11141)
/feature="match: STS: Em:HSC32F11"
10906. 11100
/feature="L2 repeat: matches 2483. .2698 of consensus"
11355. 11661
/feature="AluSg repeat: matches 2. .308 of consensus"
12351. 12478
/feature="FLAM_C repeat: matches 5. .133 of consensus"
12682. 12747
/feature="MIR repeat: matches 156. .220 of consensus"
12959. 13227
/feature="AluJo repeat: matches 1. .269 of consensus"
13293. 13459
/feature="MLTIF repeat: matches 369. .534 of consensus"
13544. 13740
/feature="MLTIF repeat: matches 87. .284 of consensus"
complement(13746. 14141)
/feature="match: GSS: Em:AQ383008"
13755. 14080
/feature="MER51A repeat: matches 50. .371 of consensus"
14141. 14179
/feature="MER51A repeat: matches 1. .39 of consensus"
14242. 14290
/feature="MLTIF repeat: matches 33. .81 of consensus"
14462. 14768
/feature="AluSg repeat: matches 1. .307 of consensus"
14890. 15196
/feature="AluSg repeat: matches 1. .302 of consensus"
15237. 15530
/feature="AluSg repeat: matches 3. .297 of consensus"
15570. 15709
/feature="MER3 repeat: matches 26. .173 of consensus"
15712. 16101
/feature="MER3 repeat: matches 2168. .2585 of consensus"
16485. 16539
/feature="MIR repeat: matches 106. .148 of consensus"
16540. 16837
```

repeat\_region /note="AluSg repeat: matches 1. .298 of consensus"  
16838. .16876  
repeat\_region /note="MIR repeat: matches 64. .106 of consensus"  
17090. .17409  
repeat\_region /note="AluSg repeat: matches 1. .312 of consensus"  
17596. .17673  
repeat\_region /note="L2 repeat: matches 2678. .2749 of consensus"  
17799. .18112  
repeat\_region /note="MLT1A1 repeat: matches 1. .365 of consensus"  
18186. .18563  
repeat\_region /note="THE1C repeat: matches 1. .371 of consensus"  
18564. .20138  
/notes="THE1C-internal repeat: matches 1. .1580 of  
consensus"  
20139. .20462  
repeat\_region /note="THE1C repeat: matches 1. .314 of consensus"  
20473. .20779  
repeat\_region /note="AluSg repeat: matches 1. .309 of consensus"  
20780. .20955  
repeat\_region /note="L1MB4 repeat: matches 6005. .6185 of consensus"  
20984. .21286  
repeat\_region /note="AluSx repeat: matches 1. .305 of consensus"  
21681. .21940  
repeat\_region /note="L1MC5 repeat: matches 7645. .7913 of consensus"  
22012. .22301  
repeat\_region /note="AluSx repeat: matches 1. .290 of consensus"  
22669. .22750  
repeat\_region /note="MER2 repeat: matches 258. .335 of consensus"  
43.5%; Score 362.6; DB 9; Length 140717;  
Best Local Similarity 70.9%; Pred. No. 7.5e-88;  
Matches 564; Conservative 0; Mismatches 199; Indels 32; Gaps 5;  
QY 48 GACACATGACCCACCGCTTGATGACAGACACCCCTCGTGTGCTTCCCA-----GTTCCAG 101  
DB 40466 GACCTCGGACGACGCGCTTGATGACAGCATCCGCGCTCTCTTTTCATCATCTCAGGGT 40525  
QY 102 TTCGGGACGCGCTCAGTCAGCGGCTCTCGCAGATAACCAAAAGCCCTGTATATCAGCAAT 161  
DB 40526 GTCCAGCAGCGCTCCATCTACAGCTTCTCCCAATACCAAGAGCTGTGTTCTCAGCAAT 40585  
QY 162 GGTGTGCGCGCAACAAAGCTCATCTGTCTAGCAACCAAGATCACCAGTACGATCAAT 221  
DB 40586 GGTGTGCGCGCAACAAAGCTCATCTGTCTAGCAACCAAGATCACCAGTACGATCAAT 40645  
QY 222 GTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCAGTACATGACGATTCAC 281  
DB 40646 GCTTCGTGGAAGTGGTCAACGATTTCTCGAGGGCATTCAGTACATAAAGGTCCTGT 40705  
QY 282 GTCATCTCCCTAACTCAGCTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 341  
DB 40706 ACCGATGCTGTGACTCGGCTCTACGACTTTTTCAGCCCATGCTGATCTTATCCAC 40765  
QY 342 AGCGTGAGATGAAGCAGGCGGCTACTTTGCTGACCTGTGCTGCTGTGAGCGGCTCA 401  
DB 40766 ACCATCATATGAGCAGGCGGCTACGCTGTGCTGACCTGTGCTGAGTGAGCGGTTCC 40825  
QY 402 GCTGCCCTGTGCTCGCTACCTCATGAAGTACCACGCCATGCTCCTGCTGGAGCGCCAC 461  
DB 40826 GCTTCAGTGTGCTGTGCTACCTCATGAATACCACTTCCATGTGCTGCTGGAGCGCCAT 40885  
QY 462 ACCTGGACCAAGTATCATCGCGGCCCATCATCCGAGCCCAACAGCGGCTTTTGGGAGCAGCTC 521  
DB 40886 ACATGGACCAAGTATCGCGGCCCATCATCCGAGCCCAACAGCGGCTTTTGGGAGCAGCTC 40945  
QY 522 ATCCATATGATGTTCCAAATGTTTGGCAAGAACACTGTGACATGTTGAGTTCCTCCAGTG 581  
DB 40946 ATCAATATGAAATTCAGAGCTGTTTAAATAAACAACACCGTGGCGCATCAACTCGCCGGA 41005  
QY 582 GGATGATCCCTGACATCTATGAGAAGAGTCCGTTTGTATGATTCACATGCTGAGCCATC 641  
DB 41006 GGTAACATCCCTGACATCTATGAGAAGAGCTAGGTATGATGATATCAATGTAGAGCCATC 41065

QY 642 CCACGACCCCTGCATTTGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGATC 701  
DB 41066 CC-----GGCCAGCCCTGCATCTGCATCTGATCTTGCACCAAGA-C 41107  
QY 702 CAACCTTGAACATCTACTTTTGTGATACAGAAAAAACAAGATGATGCCTTTTATGAGC 761  
DB 41108 TGAACCTTGAACACTGACATTTTGTAGTAAG--AAAACCGGATGGTGCCTTGTAAAGG 41165  
QY 762 AAAAAAAGA-----GTTGCTGTAGCTTTTAACTTTATATATCATTTTTCAGATT 815  
DB 41166 GCAAGAAAAAGGAGGGGTTGGAGTTTGAACGTAGTAGCCCTTACCTTAATAGATT 41225  
QY 816 AACTAATTGTGAGAT 830  
DB 41226 AATTATGAACAT 41240  
RESULT 15  
LOCUS AX040320 935 bp DNA linear PAT 18-NOV-2000  
DEFINITION Sequence 1 from Patent WO063393.  
ACCESSION AX040320  
VERSION AX040320.1 GI:11230164  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 935)  
AUTHORS Luche, R.M. and Wei, B.  
TITLE Dsp-8 dual-specificity map kinase phosphatase  
JOURNAL Patent: WO 0063393-A 1 26-OCT-2000;  
Ceptyr, Inc. (US)  
FEATURES  
Location/Qualifiers  
source  
1. .935  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 242 a 253 c 221 g 219 t  
ORIGIN  
Query Match 43.5%; Score 362.4; DB 6; Length 935;  
Best Local Similarity 73.1%; Pred. No. 3.3e-88;  
Matches 481; Conservative 0; Mismatches 171; Indels 6; Gaps 1;  
QY 48 GACACATGACCCACCGCTTGATGACAGACACCCCTCGTGTGCTTCCCA-----GTTCCAG 101  
DB 156 GACCTCGGACGACGCGCTTGATGACAGCATCCGCGCTCTCTTTTCATCATCTCAGGGT 215  
QY 102 TTCGGGACGCGCTCAGTCAGCGGCTCTCGCAGATAACCAAAAGCCCTGTATATCAGCAAT 161  
DB 216 GTCCAGCAGCGCTCCATCTACAGCTTCTCCCAATACCAAGAGCTGTGTTCTCAGCAAT 275  
QY 162 GGTGTGCGCGCAACAAAGCTCATCTGTCTAGCAACCAAGATCACCAGTACGATCAAT 221  
DB 276 GGTGTGCGCGCAACAAAGCTCATCTGTCTAGCAACCAAGATCACCAGTACGATCAAT 335  
QY 222 GTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCAGTACATGACGATTCAC 281  
DB 336 GCTTCGTGGAAGTGGTCAACGATTTCTCGAGGGCATTCAGTACATAAAGGTCCTGT 395  
QY 282 GCTGACTCCCTAACTCAGCTCTCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAC 341  
DB 396 ACCGATGCTGTGACTCGGCTCTACGACTTTTTCAGCCCATGCTGATCTTATCCAC 455  
QY 342 AGCGTGAGATGAAGCAGGCGGCTACTTTGCTGACCTGTGCTGCTGTGAGCGGCTCA 401  
DB 456 ACCATCATATGAGCAGGCGGCTACGCTGTGCTGACCTGTGCTGAGTGAGCGGTTCC 515  
QY 402 GCTGCCCTGTGCTCGCTACCTCATGAAGTACCACGCCATGCTCCTGCTGGAGCGCCAC 461  
DB 516 GCTTCAGTGTGCTGTGCTACCTCATGAATACCACTTCCATGCTGCTGAGCGCCAT 575  
QY 462 ACCTGGACCAAGTATCATCGCGGCCCATCATCCGAGCCCAACAGCGGCTTTTGGGAGCAGCTC 521

|    |     |                                                               |     |
|----|-----|---------------------------------------------------------------|-----|
| Db | 576 | ACATGGACCAAGTCGCGCCGCCCATCATCCGGCCCAACAACGCGTTTTGGGAACAGCTC   | 635 |
| Qy | 522 | ATCCACTATGAGTTCGAATGTTTGGCAAGAACACTGTGCACATGGTCACTTCCCCAGTG   | 581 |
| Db | 636 | ATCAATTACGAATTCAGCTGTTTAATAACAACACCCGTCGCATGATCAACTCGCCGTA    | 695 |
| Qy | 582 | GGAATGATCCCTGACATCTATGAGAAGGAAAGTCCGTTTGATGATTCGACTGTGAGCCATC | 641 |
| Db | 696 | GGTAACATCCCTGACATCTATGAGAGGACCTAGTATGATGATATCAATGTAGCCATC     | 755 |
| Qy | 642 | CCACGAGCCCTGCAATGGAGTCAGAGGTACAGATCTATTGTTGATCTTACACCAAGA     | 699 |
| Db | 756 | CCGSCCAGCCCTGACATCTGCCATCGATCTTGACCAAGACTGAACCTTGAACACTGA     | 813 |

Search completed: April 21, 2003, 17:45:07  
Job time : 2384 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:40:53 ; Search time 184 Seconds  
(without alignments)  
10207.427 Million cell updates/sec

Title: US-09-527-376-1

Perfect score: 834

Sequence: 1 ctttctctgtatttttgc.....aaactaattgtgagatggtg 834

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 834   | 100.0       | 834    | 21 | AAA96619    |
| 2          | 805.6 | 96.6        | 1140   | 22 | AAS35372    |
| 3          | 805.6 | 96.6        | 1140   | 22 | AAS35019    |
| 4          | 805.6 | 96.6        | 1394   | 22 | AAI58479    |
| 5          | 805.6 | 96.6        | 1407   | 22 | AAH22963    |
| 6          | 804   | 96.4        | 7537   | 22 | AAS36235    |
| 7          | 804   | 96.4        | 7537   | 22 | AAK65643    |
| 8          | 793   | 95.1        | 1783   | 22 | AAI60265    |
| 9          | 704.6 | 84.5        | 775    | 22 | AAF63573    |

|   |    |       |      |       |    |          |
|---|----|-------|------|-------|----|----------|
| c | 10 | 496   | 59.5 | 496   | 22 | ABA48128 |
| c | 11 | 496   | 59.5 | 496   | 22 | ABA66007 |
| c | 12 | 496   | 59.5 | 496   | 22 | AAK14429 |
| c | 13 | 496   | 59.5 | 496   | 22 | AAK40164 |
| c | 14 | 496   | 59.5 | 496   | 22 | AAI20936 |
| c | 15 | 496   | 59.5 | 496   | 22 | AAI46179 |
| c | 16 | 496   | 59.5 | 496   | 22 | AAI06647 |
| c | 17 | 496   | 59.5 | 496   | 24 | ABS14218 |
| c | 18 | 496   | 59.5 | 496   | 22 | ABA42999 |
| c | 19 | 463   | 55.5 | 463   | 22 | ABA53415 |
| c | 20 | 463   | 55.5 | 463   | 22 | AAK01684 |
| c | 21 | 463   | 55.5 | 463   | 22 | AAK27133 |
| c | 22 | 463   | 55.5 | 463   | 22 | AAI11722 |
| c | 23 | 463   | 55.5 | 463   | 22 | AAI33028 |
| c | 24 | 463   | 55.5 | 463   | 22 | AAI01650 |
| c | 25 | 463   | 55.5 | 463   | 24 | ABS01688 |
| c | 26 | 370.4 | 44.4 | 908   | 22 | AAF63572 |
| c | 27 | 362.4 | 43.5 | 935   | 21 | AAC67878 |
| c | 28 | 361   | 43.3 | 969   | 22 | AAK52001 |
| c | 29 | 311.2 | 37.3 | 894   | 22 | AAK52985 |
| c | 30 | 178.6 | 21.4 | 1466  | 20 | AAK80831 |
| c | 31 | 177.6 | 21.3 | 1137  | 21 | AAH88550 |
| c | 32 | 177.6 | 21.3 | 1729  | 20 | AAZ28369 |
| c | 33 | 177.6 | 21.3 | 24000 | 21 | AAH8551  |
| c | 34 | 171.4 | 20.6 | 1406  | 21 | AAZ52479 |
| c | 35 | 152.6 | 18.3 | 456   | 22 | AAK43861 |
| c | 36 | 152.6 | 18.3 | 456   | 24 | ABS18090 |
| c | 37 | 123.6 | 14.8 | 347   | 22 | AAS34897 |
| c | 38 | 98.4  | 11.8 | 467   | 22 | AAK32630 |
| c | 39 | 98.4  | 11.8 | 467   | 24 | ABS07421 |
| c | 40 | 88.8  | 10.6 | 1933  | 24 | ABI99789 |
| c | 41 | 88    | 10.6 | 2000  | 24 | ABK84633 |
| c | 42 | 88    | 10.6 | 2000  | 24 | ABN97279 |
| c | 43 | 88    | 10.6 | 2000  | 24 | ABL69410 |
| c | 44 | 88    | 10.6 | 2000  | 24 | ABK35490 |
| c | 45 | 87    | 10.4 | 1209  | 24 | ABK39796 |

ALIGNMENTS

RESULT 1

AAA96619

ID AAA96619 standard; cDNA; 834 BP.

XX AAA96619;

AC AAA96619;

DT 08-FEB-2001 (first entry)

XX cDNA encoding a human dual-specificity phosphatase 2 (DSP-2).

XX Human; mitogen-activated protein kinase; MAP kinase; cell survival;  
KW dual-specificity phosphatase; DSP-2; cell proliferation; DSP-2;  
KW cell differentiation; cancer; graft-versus-host disease; allergy;  
KW autoimmune disease; metabolic disease; cell cycle abnormality; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 69...635

XX FT /\*tag= a

XX FT /product= "dual-specificity phosphatase 2 (DSP-2)"

XX WO200056899-A1.

XX 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US07589.

XX 24-MAR-1999; 990US-0125957.

XX 16-MAR-2000; 2000US-0527376.

XX (CEPT-) CEPTYR INC.

XX Luche RM, Wei B;  
XX WPI: 2000-638264/61.  
DR P-PSDB; AAB19008.  
XX  
XX Novel dual-specificity mitogen activated protein (MAP) kinase  
PT phosphatase polypeptide useful in assays to identify agents that  
PT modulate the protein activity, which are used to treat cancer and  
PT autoimmune diseases  
XX  
XX Claim 7; Fig 1: 51pp; English.  
PS  
PS The present sequence encodes a human dual-specificity mitogen-  
CC activated protein (MAP) kinase phosphatase (DSP-2). The protein has  
CC the ability to dephosphorylate an activated MAP-kinase. The DSP-2  
CC protein is used for treatment of conditions associated with cell  
CC proliferation, cell differentiation, and cell survival. The DSP-2  
CC protein is useful for screening for modulators. Agents that modulate  
CC DSP-2 activity are useful for treating cancer, graft-versus-host  
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal  
CC cell growth, abnormal cell proliferation and cell cycle abnormalities.  
XX  
XX Sequence 834 BP; 197 A; 226 C; 183 G; 228 T; 0 other;  
SQ  
Query Match 100.0%; Score 834; DB 21; Length 834;  
Best Local Similarity 100.0%; Pred. No. 1e-237;  
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CTTTTCCTGTATTTTTCCTTCATCTTGTGTTTCGTCGACGTCGACCACTGACCCA 60  
Db 1 CTTTTCCTGTATTTTTCCTTCATCTTGTGTTTCGTCGACGTCGACCACTGACCCA 60  
Qy 61 CCGCCTTGATGACAGCAGCCCTGCTGCTCCAGTTCAGTTCGCGCAGCCCTCAGTCA 120  
Db 61 CCGCCTTGATGACAGCAGCCCTGCTGCTCCAGTTCAGTTCGCGCAGCCCTCAGTCA 120  
Qy 121 CGCGCCTCTCGCAGATACCAAAAGCCCTGTATATCAGCAATGGTGGCCGCCCAACA 180  
Db 121 CGCGCCTCTCGCAGATACCAAAAGCCCTGTATATCAGCAATGGTGGCCGCCCAACA 180  
Qy 181 AGCTCATGCTCTGAGCAACAGATACCATGGTTCATCAATGTCTCAGTGAGGTAGTGA 240  
Db 181 AGCTCATGCTCTGAGCAACAGATACCATGGTTCATCAATGTCTCAGTGAGGTAGTGA 240  
Qy 241 ACACCTTGTATGAGATATCCAGTACATGACAGTACCTGTGGCTGACCTCCCTTAACCTAC 300  
Db 241 ACACCTTGTATGAGATATCCAGTACATGACAGTACCTGTGGCTGACCTCCCTTAACCTAC 300  
Qy 301 GTCCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAGCGGTGGAGATGAAGCAGG 360  
Db 301 GTCCTGTGACTTCTTTGACCCCTATTGCTGACCATATCCAGCGGTGGAGATGAAGCAGG 360  
Qy 361 GCCGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 361 GCCGTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Qy 421 ACCTCATGATGACAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db 421 ACCTCATGATGACAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Qy 481 GGGCCATCATCGGACCAACAGCGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCAAAT 540  
Db 481 GGGCCATCATCGGACCAACAGCGGCTTTTGGGAGCAGCTCATCCACTATGAGTTCCAAAT 540  
Qy 541 TGTTTGGCAAGAACTGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 541 TGTTTGGCAAGAACTGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Qy 601 ATGAGAAGGAGTCCGTTTGTATGATTCACCTGTGAGCCATCCACGAGCCCTGCTGATGG 660  
Db 601 ATGAGAAGGAGTCCGTTTGTATGATTCACCTGTGAGCCATCCACGAGCCCTGCTGATGG 660

Qy 661 AGTCAGAGGTACAGATCTATTGTTGATCTTTACACCAAGATCCAAACTTGAACATCTTACT 720  
Db 661 AGTCAGAGGTACAGATCTATTGTTGATCTTTACACCAAGATCCAAACTTGAACATCTTACT 720  
Qy 721 TTTGTTGATACAGAAAAACAGATGATGCTTTTATGACACAAAAAGAGTTGCTGTA 780  
Db 721 TTTGTTGATACAGAAAAACAGATGATGCTTTTATGACACAAAAAGAGTTGCTGTA 780  
Qy 781 GCCTTTAACTTTTATATCCATTTTTCAGATTAAACTAATTTGAGATGGTG 834  
Db 781 GCCTTTAACTTTTATATCCATTTTTCAGATTAAACTAATTTGAGATGGTG 834  
RESULT 2  
AAS35372  
ID AAS35372 standard; CDNA; 1140 BP.  
XX AAS35372;  
AC AAS35372;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Human cardiovascular system antigen CDNA polynucleotide SEQ ID No 257.  
XX  
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
XX  
XX Homo sapiens.  
OS  
XX WO200155321-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01340.  
PF  
XX 31-JAN-2000; 2000US-0179055.  
XX 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205151.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 02-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 03-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-451930/48.  
XX P-PSDB; AAU22098.  
DR  
XX New cardiovascular system related polynucleotides and polypeptides,  
PT useful for diagnosing, treating and/or preventing disorders of the  
PT cardiovascular system -  
XX  
XX Claim 1; SEQ ID No 257; 674pp; English.  
XX  
CC Sequences AAS35126-AAS35740 represent cDNA molecules, which encode the  
CC cardiovascular system antigen polypeptides of the invention.  
CC Cardiovascular system antigens and their associated polynucleotides are  
CC useful in the diagnosis, treatment and prevention of various types of  
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. A pathological condition can be determined by  
CC detecting the presence or absence of a mutation in a cardiovascular  
CC system antigen polynucleotide. The treatable disorders include autoimmune  
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such  
CC as neoplasms of the breast or liver, cardiovascular disorders such as  
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,  
CC nervous system disorders such as Alzheimer's disease, infections caused  
CC by bacteria, viruses and fungi, ocular disorders such as corneal  
CC infection, endocrine disorders such as premature labour and infertility,  
CC gastrointestinal disorders such as Crohn's disease, renal disorders such  
CC as glomerulonephritis and respiratory disorders such as asthma and  
CC pleurisy. The polypeptides can also be used to aid wound healing, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, to regenerate tissues and in chemotaxis.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

Query Match 96.6%; Score 805.6; DB 22; Length 1140;  
Best Local Similarity 99.5%; Pred. No. 3.4e-229;  
Matches 808; Conservative 0; Mismatches 4; Indels 0; Gaps 0;





Db 696 CGGCTTTGGGAGCAGCTATCCACTATGAGTTCCAAATGTTTGGCAAGAACACTGTGCA 755  
 Qy 563 CATGGTCAGTCCCGAGTGGGAATGATCCCTGACATCTATCAGAAGGAAGTCCGTTTGAT 622  
 Db 756 CATGGTCAGTCCCGAGTGGGAATGATCCCTGACATCTATCAGAAGGAAGTCCGTTTGAT 815  
 Qy 623 GATTCCTAGTGGAGCCATCCACGAGCCCTGCGATGCGAGTGCAGAGTACAGATCTATTG 682  
 Db 816 GATTCCTAGTGGAGCCATCCACGAGCCCTGCGATGCGAGTGCAGAGTACAGATCTATTG 875  
 Qy 683 TTGATCTTACACCAAGATCCAACTTCAACATCTACTTTTGTGTACAGAAAAAACA 742  
 Db 876 TTGATCTTACACCAAGATCCAACTTCAACATCTACTTTTGTGTACAGAAAAAACA 935  
 Qy 743 GATGATGCCCTTTTATGAGCACAAGAGTTGCTGTAGCTTTTAACTTTTATATCCATT 802  
 Db 936 GATGATGCCCTTTTATGAGCACAAGAGTTGCTGTAGCTTTTAACTTTTATATCCATT 995  
 Qy 803 TTTTTCAGATTAACATAATGTGAGATGGTG 834  
 Db 996 TTTTTCAGATTAACATAATGTGAGATGGTG 1027

## RESULT 4

AA158479  
 ID AA158479 standard; cDNA; 1394 BP.

XX AC AA158479;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 682.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0820312.

XX PR 03-AUG-2000; 2000US-0853450.

XX PR 14-SEP-2000; 2000US-0862191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX DR WPI: 2001-442253/47.

XX DR P-PSDB: AAM39323.

XX XX Novel nucleic acids and polypeptides, useful for treating disorders  
 XX PT such as central nervous system injuries -  
 XX PS Claim 1; SEQ ID NO 682; 10078pp; English.  
 XX XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AA42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX SQ Sequence 1394 BP; 315 A; 385 C; 377 G; 317 T; 0 other;

Query Match 96.8%; Score 805.6; DB 22; Length 1394;  
 Best Local Similarity 99.5%; Pred. No. 3.9e-229;  
 Matches 808; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 23 CATCTTGGTGTTCGCTGACTGCTGACCACTGACCCACCCGCTTGTGATGACAGCACCCCTC 82

Db 441 CAGCCTTCGTGCTGCTGACTGCTGACCACTGACCCACCCGCTTGTGATGACAGCACCCCTC 500

Qy 83 GTGTGCTTCCCAAGTTTCAGTTCGGGAGCCCTCAGTCAGCGGGCTCTCGCAGATAACCAA 142

Db 501 GTGTGCTTCCCAAGTTTCAGTTCGGGAGCCCTCAGTCAGCGGGCTCTCGCAGATAACCAA 560

Qy 143 AAGCCTGTATATCAGCAATGGTGTGGCCGCCAACACAAAGCTCATGTCTGTAGCAACCA 202

Db 561 AAGCCTGTATATCAGCAATGGTGTGGCCGCCAACACAAAGCTCATGTCTGTAGCAACCA 620

Qy 203 GATCACCATGGTCATCAATGCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 262

Db 621 GATCACCATGGTCATCAATGCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 680

Qy 263 GTACATCAGGTACCTGTGGCTGACTCCCTCACTCACTCACTCTCTGTGACTTCTTTGACCC 322

Db 681 GTACATCAGGTACCTGTGGCTGACTCCCTCACTCACTCACTCTCTGTGACTTCTTTGACCC 740

Qy 323 TATTGTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCGTACTTTGTGCTGCTGTC 382

Db 741 TATTGTGACCATATCCACAGCGTGGAGATGAAGCAGGCGCGTACTTTGTGCTGCTGTC 800

Qy 383 TGCTGTGTGAGCGGCTCAGCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442

Db 801 TGCTGTGTGAGCGGCTCAGCTGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860

Qy 443 GTCCTGCTGGAGCGCCACACAGTGGACCAAGTCAATGCGGCCCATCATCGACCCAAACAG 502

Db 861 GTCCTGCTGGAGCGCCACACAGTGGACCAAGTCAATGCGGCCCATCATCGACCCAAACAG 920

Qy 503 CGGCTTTGGGAGCAGCTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACACTGTGCA 562

Db 921 CGGCTTTGGGAGCAGCTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACACTGTGCA 980

Qy 563 CATGGTCAGTTCGCCAGTGGGAATGATCCCTGACATCTATCAGAGGAAGTCCGTTTGAT 622

Db 981 CATGGTCAGTTCGCCAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 1040

Qy 623 GATTCCACTGTAGCCCATCCACGAGCCCTGCAATGAGTTCAGAGTACAGATCTATTG 682

Db 1041 GATTCCACTGTAGCCCATCCACGAGCCCTGCAATGAGTTCAGAGTACAGATCTATTG 1100

Qy 683 TTGATCTTACCAAGATCCAACTTGAACATTTCTACTTTTGTGTGATACAGAAAAAACA 742

Db 1101 TTGATCTTACCAAGATCCAACTTGAACATTTCTACTTTTGTGTGATACAGAAAAAACA 1160

Qy 743 GATGATGCCCTTTTATGAGCACAAGAGTTGCTGTAGCTTTTAACTTTTATATCCATT 802

Db 1161 GATGATGCCCTTTTATGAGCACAAGAGTTGCTGTAGCTTTTAACTTTTATATATCCATT 1220



XX OS Homo sapiens.  
XX PN WO200155321-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01340.  
XX PR 31-JAN-2000; 2000US-0179065.  
XX PR 04-FEB-2000; 2000US-0180828.  
XX PR 24-FEB-2000; 2000US-0184664.  
XX PR 02-MAR-2000; 2000US-0186350.  
XX PR 16-MAR-2000; 2000US-0189874.  
XX PR 17-MAR-2000; 2000US-0190076.  
XX PR 18-APR-2000; 2000US-0198123.  
XX PR 19-MAY-2000; 2000US-0205515.  
XX PR 07-JUN-2000; 2000US-0209467.  
XX PR 28-JUN-2000; 2000US-0214886.  
XX PR 30-JUN-2000; 2000US-0215135.  
XX PR 07-JUL-2000; 2000US-0216647.  
XX PR 07-JUL-2000; 2000US-0216880.  
XX PR 11-JUL-2000; 2000US-0217487.  
XX PR 11-JUL-2000; 2000US-0217496.  
XX PR 14-JUL-2000; 2000US-0218290.  
XX PR 26-JUL-2000; 2000US-0220963.  
XX PR 26-JUL-2000; 2000US-0220964.  
XX PR 14-AUG-2000; 2000US-0224518.  
XX PR 14-AUG-2000; 2000US-0224519.  
XX PR 14-AUG-2000; 2000US-0225213.  
XX PR 14-AUG-2000; 2000US-0225214.  
XX PR 14-AUG-2000; 2000US-0225266.  
XX PR 14-AUG-2000; 2000US-0225267.  
XX PR 14-AUG-2000; 2000US-0225268.  
XX PR 14-AUG-2000; 2000US-0225270.  
XX PR 14-AUG-2000; 2000US-0225447.  
XX PR 14-AUG-2000; 2000US-0225757.  
XX PR 14-AUG-2000; 2000US-0225758.  
XX PR 14-AUG-2000; 2000US-0225759.  
XX PR 18-AUG-2000; 2000US-0226279.  
XX PR 22-AUG-2000; 2000US-0226681.  
XX PR 22-AUG-2000; 2000US-0226868.  
XX PR 22-AUG-2000; 2000US-0227182.  
XX PR 23-AUG-2000; 2000US-0227009.  
XX PR 30-AUG-2000; 2000US-0228924.  
XX PR 01-SEP-2000; 2000US-0229287.  
XX PR 01-SEP-2000; 2000US-0229343.  
XX PR 01-SEP-2000; 2000US-0229344.  
XX PR 01-SEP-2000; 2000US-0229345.  
XX PR 05-SEP-2000; 2000US-0229509.  
XX PR 05-SEP-2000; 2000US-0229513.  
XX PR 06-SEP-2000; 2000US-0230437.  
XX PR 06-SEP-2000; 2000US-0230438.  
XX PR 08-SEP-2000; 2000US-0231242.  
XX PR 08-SEP-2000; 2000US-0231243.  
XX PR 08-SEP-2000; 2000US-0231244.  
XX PR 08-SEP-2000; 2000US-0231413.  
XX PR 08-SEP-2000; 2000US-0231414.  
XX PR 08-SEP-2000; 2000US-0232080.  
XX PR 08-SEP-2000; 2000US-0232081.  
XX PR 12-SEP-2000; 2000US-0231968.  
XX PR 14-SEP-2000; 2000US-0232397.  
XX PR 14-SEP-2000; 2000US-0232398.  
XX PR 14-SEP-2000; 2000US-0232399.  
XX PR 14-SEP-2000; 2000US-0232399.  
XX PR 14-SEP-2000; 2000US-0232400.  
XX PR 14-SEP-2000; 2000US-0232401.  
XX PR 14-SEP-2000; 2000US-0233063.  
XX PR 14-SEP-2000; 2000US-0233064.  
XX PR 21-SEP-2000; 2000US-0233065.  
XX PR 21-SEP-2000; 2000US-0234223.  
XX PR 21-SEP-2000; 2000US-0234274.  
XX PR 25-SEP-2000; 2000US-0234397.  
XX PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249266.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2001US-0259678.  
XX XX







|    |                                                                           |  |
|----|---------------------------------------------------------------------------|--|
| OS | Homo sapiens.                                                             |  |
| XX |                                                                           |  |
| PN | WO200153312-A1.                                                           |  |
| XX |                                                                           |  |
| XX |                                                                           |  |
| PD | 26-JUL-2001.                                                              |  |
| XX |                                                                           |  |
| XX | 26-DEC-2000; 2000WO-US34263.                                              |  |
| PF |                                                                           |  |
| XX |                                                                           |  |
| XX | 21-JAN-2000; 2000US-0488725.                                              |  |
| PR | 25-APR-2000; 2000US-0552317.                                              |  |
| PR | 03-JUL-2000; 2000US-0598042.                                              |  |
| PR | 19-JUL-2000; 2000US-0620312.                                              |  |
| PR | 03-AUG-2000; 2000US-0653450.                                              |  |
| PR | 14-SEP-2000; 2000US-0662191.                                              |  |
| PR | 19-OCT-2000; 2000US-0593036.                                              |  |
| PR | 29-NOV-2000; 2000US-0727344.                                              |  |
| XX |                                                                           |  |
|    | (HYSE-) HYSEQ INC.                                                        |  |
| XX |                                                                           |  |
| XX | Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;           |  |
| PI | Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;                 |  |
| PI | Zhao QA, Zhou P, Goodrich R, Drmanac RT;                                  |  |
| XX |                                                                           |  |
| XX | WPI; 2001-442253/47.                                                      |  |
| DR | P-PSDB; AAM41109.                                                         |  |
| DR |                                                                           |  |
| XX |                                                                           |  |
| PT | Novel nucleic acids and polypeptides, useful for treating disorders       |  |
| PT | such as central nervous system injuries -                                 |  |
| XX |                                                                           |  |
| XX | Claim 1; SEQ ID NO 4254; 10078pp; English.                                |  |
| XX |                                                                           |  |
| CC | The invention relates to human nucleic acids (AAI57798-AAI61369) and      |  |
| CC | the encoded polypeptides (AAM38642-AAM42213) with nootropic,              |  |
| CC | immunosuppressant and cytostatic activity. The polynucleotides are useful |  |
| CC | in gene therapy. A composition containing a polypeptide or polynucleotide |  |
| CC | of the invention may be used to treat diseases of the peripheral nervous  |  |
| CC | system, such as peripheral nervous injuries, peripheral neuropathy and    |  |
| CC | localised neuropathies and central nervous system diseases, such as       |  |
| CC | Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic       |  |
| CC | lateral sclerosis, and Shy-Drager Syndrome. Other uses include the        |  |
| CC | utilisation of the activities such as: Immune system suppression,         |  |
| CC | Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  |  |
| CC | and thrombolytic activity, cancer diagnosis and therapy, drug screening,  |  |
| CC | assays for receptor activity, arthritis and inflammation, leukaemias and  |  |
| CC | C.N.S disorders.                                                          |  |
| CC | Note: The sequence data for this patent did not form part of the printed  |  |
| CC | specification.                                                            |  |
| XX |                                                                           |  |
| SQ | Sequence 1783 BP; 405 A; 477 C; 457 G; 436 T; 8 other;                    |  |
|    |                                                                           |  |
|    | Query Match 95.1%; Score 793; DB 22; Length 1783;                         |  |
|    | Best Local Similarity 99.3%; Pred. No. 2.5e-225;                          |  |
|    | Matches 807; Conservative 0; Mismatches 5; Indels 1; Gaps 1;              |  |

|    |  |     |                                                                  |     |
|----|--|-----|------------------------------------------------------------------|-----|
| Db |  | 441 | CAGCCTTCGTGCTTCGGTGAATGCTGACCACTGACCAACCACCCCTTGATGACAGACCCCTC   | 500 |
| Qy |  | 83  | GTGTGCCCTTCCCAGTTTCAGTTTCGGGCAGCCCTCAGTCAGCGGCCCTCTCGCAGATAACCAA | 142 |
| Db |  | 501 | GTGTGCCCTTCCCAGTTTCAGTTTCGGGCAGCCCTCAGTCAGCGGCCCTCTCGCAGATAACCAA | 560 |
| Qy |  | 143 | AAGCCTCTATATACAGCAATGGTGTGGCGCGCAACAAGAAGTCATGCTGTAGCAACCA       | 202 |
| Db |  | 561 | AAGCCTCTATATACAGCAATGGTGTGGCGCGCAACAAGAAGTCATGCTGTAGCAACCA       | 620 |
| Qy |  | 203 | GATCACCATGGTCATCAATGTCCTCAGTGGAGGTAGTGAACACCTTGATGAGGATATCCA     | 262 |
| Db |  | 621 | GATCACCATGGTCATCAATGTCCTCAGTGGAGGTAGTGAACACCTTGATGAGGATATCCA     | 680 |
| Qy |  | 263 | GTACATCAGGTAACCTGTGGCTGACTCCCCCTAACTACGTCCTGTGACTTCTTTGACCC      | 322 |
| Db |  | 681 | GTACATCAGGTAACCTGTGGCTGACTCCCCCTAACTACGTCCTGTGACTTCTTTGACCC      | 740 |

QY 323 TATTGCTGACCATATCCACAGCGGTGAGATGAAGCAGGCGG-TACTTTGCTGCACTGTG 381  
DB 741 TATTGCTGACCATATCCACAGCGGTGAGATGAAGCAGGCGGTACTTTGCTGCACTGTG 800  
QY 382 CTGCTGTGTGAGCGGTCTAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCACGCCA 441  
DB 801 CTGCTGTGTGAGCGGTCTAGCTGCCCTGTGCTGCCCTACCTCATGAAGTACCACGCCA 860  
QY 442 TGTCCCTGCTGGAGCGCCACAGCTGGACCAAGTCATCCCGGCCCATATCCGACCCCAACA 501  
DB 861 TGTCCCTGCTGGAGCGCCACAGCTGGACCAAGTCATCCCGGCCCATATCCGACCCCAACA 920  
QY 502 GCGGCTTTTGGGACGAGCTATCCACATGAGTTTCCAAATTTTGGCAAGAACACTGTGC 561  
DB 921 GCGGCTTTTGGGACGAGCTATCCACATGAGTTTCCAAATTTTGGCAAGAACACTGTGC 980  
QY 562 ACATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAAGAAAGTCCGTTTGA 621  
DB 981 ACATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAAGAAAGTCCGTTTGA 1040  
QY 622 TGATTCACATCTGAGCCATCCACAGCGCCCTGCATTTGGAGTCAGAGTACAGATCTATT 681  
DB 1041 TGATTCACATCTGAGCCATCCACAGCGCCCTGCATTTGGAGTCAGAGTACAGATCTATT 1100  
QY 682 GTTGATCTTACACCAAGATCCAACTTTGAACATTTCTACTTTTGTGATACAGAAAAAACC 741  
DB 1101 GTTGATCTTACACCAAGATCCAACTTTGAACATTTCTACTTTTGTGATACAGAAAAAACC 1160  
QY 742 AGATGATGCCCTTTATGAGCACAAGAGTGTCTGTAGCTTTTAACTTTTAAATCCAT 801  
DB 1161 AGATGATGCCCTTTATGAGCACAAGAGTGTCTGTAGCTTTTAACTTTTAAATCCAT 1220  
QY 802 TTTTTCAGATTAACATTAATTGTGAGATGGTG 834  
DB 1221 TTTTTCAGATTAACATTAATTGTGAGATGGTG 1253

RESULT 9

ID AAF63573 standard; cDNA; 775 BP.

XX AAF63573;

DT 11-MAY-2001 (first entry)

XX Human phosphatase AA915932\_h coding sequence.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;  
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;  
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;  
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;  
KW schizophrenia; hamartoma; ss.

XX Homo sapiens.

OS WO200112819-A2.

PN 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

XX 13-AUG-1999; 99US-0149005.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

XX WPI; 2001-211226/21.

XX P-PSDB; AAB73221.

XX New protein phosphatase polypeptide for diagnosing and treating  
PT phosphatase related disorders such as cancer, schizophrenia, cardiac

dysfunction and/or vascular disorders

Claim 1; Fig 4; 138pp; English.

CC The present invention relates to phosphatase proteins and coding  
sequences. The present sequence is one such phosphatase coding sequence.  
CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins  
CC modified by phosphorylation of serine, threonine or tyrosine residues.  
CC The phosphatases are useful for treating a variety of diseases: for  
CC example cancer e.g. breast, urogenital, prostate, head, neck, lung  
CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung  
CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach  
CC cancer, glioblastoma, colorectal cancer and thyroid cancer,  
CC pathophysiological hypoxia, cardiac dysfunction and/or vascular  
CC disorders, myopathies, congenital muscle disorders, Papillon-Lefevre  
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,  
CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and  
CC hamartomas.

XX SQ Sequence 775 BP; 179 A; 229 C; 180 G; 187 T; 0 other;

Query Match 84.5%; Score 704.6; DB 22; Length 775;

Best Local Similarity 99.4%; Pred. No. 3.2e-199;

Matches 707; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 23 CATTCCTTGTGTTTTCGCTGACTGCTGACCACTGACCCACCGCCCTTGATGACAGCACCCCTC 82

DB 65 CAGCCTTCGCTGCTGCTGACTGCTGACCACTGACCCACCGCCCTTGATGACAGCACCCCTC 124

QY 83 GTGTGCTTCCCGATTCAGTTCCGGCAGCCCTCAGTCAGCGGCCCTTCGCGAGATAACCAA 142

DB 125 GTGTGCTTCCCGATTCAGTTCCGGCAGCCCTCAGTCAGCGGCCCTTCGCGAGATAACCAA 184

QY 143 AAGCCTGTATATCAGCAATGTTGGCCGCCAACAACAGCTCATGCTGTACCAACCA 202

DB 185 AAGCCTGTATATCAGCAATGTTGGCCGCCAACAACAGCTCATGCTGTACCAACCA 244

QY 203 GATCACCATTGCTCATCAATGCTCAGTGGAGTAGTGAACACCTTGTATGAGGATATCCA 262

DB 245 GATCACCATTGCTCATCAATGCTCAGTGGAGTAGTGAACACCTTGTATGAGGATATCCA 304

QY 263 GTACATGCAAGTACCTGTGGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTGACCC 322

DB 305 GTACATGCAAGTACCTGTGGCTGACTCCCTTAACCTACAGTCTCTGTGACTTCTTTGACCC 364

QY 323 TATTGCTGACCATATCCACAGCGTGGAGATCAACAGCGCGGTACTTTGCTGCACTGTGC 382

DB 365 TATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGCGCGGTACTTTGCTGCACTGTGC 424

QY 383 TGTGCTGTGAGCGGCTCAGCTGCCCTGTGCTGCTGCTGCTATGAAAGTACCACGCCAT 442

DB 425 TGTGCTGTGAGCGGCTCAGCTGCCCTGTGCTGCTGCTATGAAAGTACCACGCCAT 484

QY 443 GTCCTCTGCTGGACGCCACACAGCTGGACCAAGTCATGCGGCCCATCATCGACCCACAG 502

DB 485 GTCCTCTGCTGGACGCCACACAGCTGGACCAAGTCATGCGGCCCATCATCGACCCACAG 544

QY 503 CGGCTTTTGGGACGAGCTCATCCACTATGATTCCTCAATTTGTTGGCAAGAACACTGTGCA 562

DB 545 CGGCTTTTGGGACGAGCTCATCCACTATGATTCCTCAATTTGTTGGCAAGAACACTGTGCA 604

QY 563 CATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGAT 622

DB 605 CATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCCTTTGAT 664

QY 623 GATTCCACTGTGAGCCCATCCACAGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682

DB 665 GATTCCACTGTGAGCCCATCCACAGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724

QY 683 TTGATCTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGATACAG 733

DB 725 TTGATCTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGATACAG 775



CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;

Query Match 59.5%; Score 496; DB 22; Length 496;  
Best Local Similarity 100.0%; Pred. No. 3.le-137;  
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 TTCGGGAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCTGTATATCAGCAAT 161  
DB 496 TTCGGGAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCTGTATATCAGCAAT 437  
QY 162 GGTGTGGCCGCCAACAAAGCTCATGCTGTAGCAACAGATCACCATTGGTCAAT 221  
DB 436 GGTGTGGCCGCCAACAAAGCTCATGCTGTAGCAACAGATCACCATTGGTCAAT 377  
QY 222 GTCTCAGTGAGGTAGTAACACCTTGTATGAGGATATCCAGTACAGGTACCTGTG 281  
DB 376 GTCTCAGTGAGGTAGTAACACCTTGTATGAGGATATCCAGTACAGGTACCTGTG 317  
QY 282 GCTGACTCCCTTAACCTACCTCTCTGTGACTTCTTTGACCTATTGCTGACCATATCCAC 341  
DB 316 GCTGACTCCCTTAACCTACCTCTCTGTGACTTCTTTGACCTATTGCTGACCATATCCAC 257  
QY 342 AGCGTGAGATGAAGCAGGCGCTACTTTGCTGCATGTGCTGCTGTGAGCGCCGCTCA 401  
DB 256 AGCGTGAGATGAAGCAGGCGCTACTTTGCTGCATGTGCTGCTGTGAGCGCCGCTCA 197  
QY 402 GCTGCGCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 461  
DB 196 GCTGCGCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 137  
QY 462 ACCTGGACCAAGTCATGCGCGCCCATCATCCGACCCAAAGCGGCTTTTGGGAGCAGCTC 521  
DB 136 ACCTGGACCAAGTCATGCGCGCCCATCATCCGACCCAAAGCGGCTTTTGGGAGCAGCTC 77  
QY 522 ATCCACTATGAGTTCCAAATTTTGGCAAGAACACTGTGCACATGCTATGTCATGTC 581  
DB 76 ATCCACTATGAGTTCCAAATTTTGGCAAGAACACTGTGCACATGCTATGTCATGTC 17  
QY 582 GGAATGATCCCTGACA 597  
DB 16 GGAATGATCCCTGACA 1

RESULT 12  
AAK1429/C  
ID AAK1429 standard; DNA; 496 BP.  
XX  
AC AAK1429;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe SEQ ID NO: 14420.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483446/52.  
DR  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT brains  
XX  
PS Example 4; SEQ ID NO: 14420; 650pp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention.  
XX  
SQ Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;

Query Match 59.5%; Score 496; DB 22; Length 496;  
Best Local Similarity 100.0%; Pred. No. 3.le-137;  
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 TTCGGGAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCTGTATATCAGCAAT 161  
DB 496 TTCGGGAGCCCTCAGTCAGCGCCCTCTCGCAGATACCAAAAGCTGTATATCAGCAAT 437  
QY 162 GGTGTGGCCGCCAACAAAGCTCATGCTGTAGCAACAGATCACCATTGGTCAAT 221  
DB 436 GGTGTGGCCGCCAACAAAGCTCATGCTGTAGCAACAGATCACCATTGGTCAAT 377  
QY 222 GTCTCAGTGAGGTAGTAACACCTTGTATGAGGATATCCAGTACAGGTACCTGTG 281  
DB 376 GTCTCAGTGAGGTAGTAACACCTTGTATGAGGATATCCAGTACAGGTACCTGTG 317  
QY 282 GCTGACTCCCTTAACCTACCTCTCTGTGACTTCTTTGACCTATTGCTGACCATATCCAC 341  
DB 316 GCTGACTCCCTTAACCTACCTCTCTGTGACTTCTTTGACCTATTGCTGACCATATCCAC 257  
QY 342 AGCGTGAGATGAAGCAGGCGCTACTTTGCTGCATGTGCTGCTGTGAGCGCCGCTCA 401  
DB 256 AGCGTGAGATGAAGCAGGCGCTACTTTGCTGCATGTGCTGCTGTGAGCGCGCTCA 197  
QY 402 GCTGCGCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 461  
DB 196 GCTGCGCTGTGCGCTACCTCATGAGTACCAAGCCATGTCCTGCTGGAGCGCCAC 137  
QY 462 ACCTGGACCAAGTCATGCGCGCCCATCATCCGACCCAAAGCGGCTTTTGGGAGCAGCTC 521  
DB 136 ACCTGGACCAAGTCATGCGCGCCCATCATCCGACCCAAAGCGGCTTTTGGGAGCAGCTC 77  
QY 522 ATCCACTATGAGTTCCAAATTTTGGCAAGAACACTGTGCACATGCTATGTCATGTC 581  
DB 76 ATCCACTATGAGTTCCAAATTTTGGCAAGAACACTGTGCACATGCTATGTCATGTC 17  
QY 582 GGAATGATCCCTGACA 597  
DB 16 GGAATGATCCCTGACA 1

RESULT 13  
AAK40164/C  
ID AAK40164 standard; DNA; 496 BP.  
XX

AAK40164;  
06-NOV-2001 (first entry)  
Human bone marrow expressed single exon probe SEQ ID NO: 14721.  
Human; bone marrow expressed exon; gene expression analysis; probe;  
microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
Homo sapiens.  
WO200157276-A2.  
09-AUG-2001.  
30-JAN-2001; 2001WO-US00668.  
04-FEB-2000; 2000US-0180312.  
26-MAY-2000; 2000US-0207456.  
30-JUN-2000; 2000US-0608408.  
03-AUG-2000; 2000US-0632366.  
21-SEP-2000; 2000US-0234687.  
27-SEP-2000; 2000US-0236359.  
04-OCT-2000; 2000GB-0024263.  
(MOLE-) MOLECULAR DYNAMICS INC.  
Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2001-488900/53.  
Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human bone marrow -  
Example 4; SEQ ID NO: 14721; 658pp + Sequence Listing; English.  
The present invention provides a number of single exon nucleic acid  
probes which are derived from genomic sequences expressed in the human  
bone marrow. They can be used to measure gene expression in bone marrow  
samples, which may enable the improved diagnosis and treatment of cancers  
such as lymphoma, leukemia and myeloma. The present sequence is one of  
the probes of the invention.  
Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;  
Query Match 59.5%; Score 496; DB 22; Length 496;  
Best Local Similarity 100.0%; Pred. No. 3.1e-137; Indels 0; Gaps 0;  
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 102 TTCGGCAGCCCTCAGTCAGCGGCGCTCTCGCAGATAACCAAAAGCCTGTATATCAGCAAT 161  
Db 496 TTCGGCAGCCCTCAGTCAGCGGCGCTCTCGCAGATAACCAAAAGCCTGTATATCAGCAAT 437  
Qy 162 GGTGTGGCGCCACACAAAGCTAAGTGTCTAGCAACCAAGATCACCATGGTCATCAAT 221  
Db 436 GGTGTGGCGCCACACAAAGCTAAGTGTCTAGCAACCAAGATCACCATGGTCATCAAT 377  
Qy 222 GTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATCAGGTACCTGTG 281  
Db 376 GTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCAGTACATCAGGTACCTGTG 317  
Qy 282 GCTGACTCCCTTAAGTACAGTCTGTGTGACTTCTTTGACCTATTGCTGACCATATCCAC 341  
Db 316 GCTGACTCCCTTAAGTACAGTCTGTGTGACTTCTTTGACCTATTGCTGACCATATCCAC 257  
Qy 342 AGCGTGGAGATGAACAGCGGCGGTACTTTGTGTGACTGTGTGTGTGAGCGGCTCA 401  
Db 256 AGCGTGGAGATGAACAGCGGCGGTACTTTGTGTGACTGTGTGTGTGAGCGGCTCA 197  
Qy 402 GCTGCGCTGTGCTGCGCTACCTCATATGAAGTACCAAGCCATGTCCCTGCTGAGCCGCAC 461  
Db 196 GCTGCGCTGTGCTGCGCTACCTCATATGAAGTACCAAGCCATGTCCCTGCTGAGCCGCAC 137

Qy 462 ACGTGGACCAAGTCATGCGGCGCCCATCATCGACCAACAGCGGCTTTTGGAGCAGCTC 521  
Db 136 ACGTGGACCAAGTCATGCGGCGCCCATCATCGACCAACAGCGGCTTTTGGAGCAGCTC 77  
Qy 522 ATCCACTATGAGTTCCCAATTTGTTGGCAAGAACACTGTGCACATGTCAGTTCCTCCAGTG 581  
Db 76 ATCCACTATGAGTTCCCAATTTGTTGGCAAGAACACTGTGCACATGTCAGTTCCTCCAGTG 17  
Qy 582 GGAATGATCCCTGACA 597  
Db 16 GGAATGATCCCTGACA 1  
RESULT 14  
AAI20936/C  
ID AAI20936 standard; DNA; 496 BP.  
XX AC AAI20936;  
DT 12-OCT-2001 (first entry)  
XX DE Probe #10869 for gene expression analysis in human cervical cell sample.  
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;  
XX KW cervical cancer; ss.  
XX OS Homo sapiens.  
XX PN WO200157278-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US00670.  
XX PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;  
PI WPI; 2001-488901/53.  
XX DR Human genome-derived single exon nucleic acid probes useful for  
XX PT analyzing gene expression in human cervical epithelial cells -  
XX PS Claim 25; SEQ ID NO 10869; 487pp; English.  
XX CC The present invention relates to human single exon nucleic acid probes  
CC (SNP). The present sequence is one such probe. The SNPs are derived  
CC from human HeLa cells. The SNPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 496 BP; 115 A; 119 C; 151 G; 111 T; 0 other;  
Query Match 59.5%; Score 496; DB 22; Length 496;  
Best Local Similarity 100.0%; Pred. No. 3.1e-137; Indels 0; Gaps 0;  
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 102 TTCGGCAGCCCTCAGTCAGCGGCGCTCTCGCAGATAACCAAAAGCCTGTATATCAGCAAT 161  
Db 496 TTCGGCAGCCCTCAGTCAGCGGCGCTCTCGCAGATAACCAAAAGCCTGTATATCAGCAAT 437

|    |     |                                                                |     |
|----|-----|----------------------------------------------------------------|-----|
| QY | 162 | GGTGTGGCGCCAAACAAGCTCATGCTGTAGCAACAGATCACCATGGTCATCAAT         | 221 |
|    |     |                                                                |     |
| Db | 436 | GGTGTGGCGCCAAACAAGCTCATGCTGTAGCAACAGATCACCATGGTCATCAAT         | 377 |
|    |     |                                                                |     |
| QY | 222 | GTCTCAGTGGAGTAGTGAACACTTGTATGAGGATATCCAGTACATGCAGTACTGTG       | 281 |
|    |     |                                                                |     |
| Db | 376 | GTCTCAGTGGAGTAGTGAACACTTGTATGAGGATATCCAGTACATGCAGTACTGTG       | 317 |
|    |     |                                                                |     |
| QY | 282 | GCTGACTCCCTPAACTCAGTCTCTGTGACTTCTTTGACCCATTGCTGACCATATCCAC     | 341 |
|    |     |                                                                |     |
| Db | 316 | GCTGACTCCCTAACTACGCTCTCTGTGACTTCTTTGACCCATTGCTGACCATATCCAC     | 257 |
|    |     |                                                                |     |
| QY | 342 | AGCGTGGAGATGAAGCAGGCGCGTACTTTGCTGCACGTGCTGCTGGTGAGCGCGTCA      | 401 |
|    |     |                                                                |     |
| Db | 256 | AGCGTGGAGATGAAGCAGGCGCGTACTTTGCTGCACGTGCTGCTGGTGAGCGCGTCA      | 197 |
|    |     |                                                                |     |
| QY | 402 | GCTGCCCTGTGCGCTACCTCATGAAGTACCCAGCGCATGTCCCTGCTGGACGCCAC       | 461 |
|    |     |                                                                |     |
| Db | 196 | GCTGCCCTGTGCGCTACCTCATGAAGTACCCAGCGCATGTCCCTGCTGGACGCCAC       | 137 |
|    |     |                                                                |     |
| QY | 462 | ACGTGGACCAAGTCATGCGCGGCCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTC | 521 |
|    |     |                                                                |     |
| Db | 136 | ACGTGGACCAAGTCATGCGCGGCCCATCATCCGACCCCAACAGCGGCTTTTGGGAGCAGCTC | 77  |
|    |     |                                                                |     |
| QY | 522 | ATCCACTATGAGTTCCAAATGTTTGGCAAGACACTGTCCACATGCTCAGTTCCTCCGAGT   | 581 |
|    |     |                                                                |     |
| Db | 76  | ATCCACTATGAGTTCCAAATGTTTGGCAAGACACTGTCCACATGCTCAGTTCCTCCGAGT   | 17  |
|    |     |                                                                |     |
| QY | 582 | GGAAATGATCCCTGACA                                              | 597 |
|    |     |                                                                |     |
| Db | 16  | GGAAATGATCCCTGACA                                              | 1   |

|            |                                                                       |
|------------|-----------------------------------------------------------------------|
| RESULT 15  |                                                                       |
| AAI46179/c |                                                                       |
| ID         | AAI46179 standard; DNA; 496 BP.                                       |
| XX         |                                                                       |
| XX         | AAI46179;                                                             |
| XX         | AC                                                                    |
| XX         |                                                                       |
| DT         | 17-OCT-2001 (first entry)                                             |
| XX         |                                                                       |
| DE         | Probe #1485 used to measure gene expression in human placenta sample. |
| XX         |                                                                       |
| DE         | Probe; microarray; human; placenta; antenatal diagnosis;              |
| KW         | genetic disorder; ss.                                                 |
| KW         |                                                                       |
| XX         |                                                                       |
| XX         | Homo sapiens.                                                         |
| OS         |                                                                       |
| XX         |                                                                       |
| XX         | WO200157272-A2.                                                       |
| PN         |                                                                       |
| XX         |                                                                       |
| PD         | 09-AUG-2001.                                                          |
| XX         |                                                                       |
| PF         | 30-JAN-2001; 2001WO-US00663.                                          |
| XX         |                                                                       |
| XX         | 04-FEB-2000; 2000US-0180312.                                          |
| XX         | 26-MAY-2000; 2000US-0207456.                                          |
| PR         | 30-JUN-2000; 2000US-0608408.                                          |
| PR         | 03-AUG-2000; 2000US-0632366.                                          |
| PR         | 21-SEP-2000; 2000US-0234687.                                          |
| PR         | 27-SEP-2000; 2000US-0236359.                                          |
| PR         | 04-OCT-2000; 2000GB-0024263.                                          |
| XX         |                                                                       |
| XX         | (MOLE-) MOLECULAR DYNAMICS INC.                                       |
| PA         |                                                                       |
| XX         |                                                                       |
| XX         | Penn SG, Hanzel DK, Chen W, Rank DR;                                  |
| PI         |                                                                       |
| XX         | WPI; 2001-488897/53.                                                  |
| DR         |                                                                       |
| XX         |                                                                       |
| XX         | Human genome-derived single exon nucleic acid probes useful for       |
| PT         | analyzing gene expression in human placenta -                         |
| XX         |                                                                       |
| PS         | Claim 25; SEQ ID No 14865; 654pp; English.                            |



| Result No. | Score | Query |      | Length | DB                  | ID    | Description         |
|------------|-------|-------|------|--------|---------------------|-------|---------------------|
|            |       | Match |      |        |                     |       |                     |
| 1          | 805.6 | 96.6  | 1140 | 9      | US-10-091-504-257   |       | Sequence 257, App   |
| 2          | 805.6 | 96.6  | 1140 | 10     | US-09-764-869-257   |       | Sequence 257, App   |
| 3          | 804   | 96.4  | 7537 | 9      | US-10-091-504-1735  |       | Sequence 1735, App  |
| 4          | 804   | 96.4  | 7537 | 10     | US-09-764-869-1735  |       | Sequence 1735, App  |
| c 5        | 496   | 59.5  | 496  | 10     | US-09-864-761-16715 |       | Sequence 26715, App |
| c 6        | 463   | 55.5  | 463  | 10     | US-09-864-761-16715 |       | Sequence 26715, App |
| c 7        | 361   | 43.3  | 923  | 12     | US-10-044-205A-1    | 10073 | Sequence 10073, App |
| 8          | 340.8 | 40.9  | 570  | 12     | US-10-044-205A-43   |       | Sequence 41, App    |
| 9          | 152.6 | 18.3  | 455  | 10     | US-09-864-761-30844 |       | Sequence 43, App    |
| 10         | 149.6 | 17.9  | 239  | 10     | US-08-783-590-8474  |       | Sequence 30844, App |
| 11         | 98.4  | 11.8  | 467  | 10     | US-09-864-761-14459 |       | Sequence 8474, App  |
| 12         | 88    | 10.6  | 2000 | 10     | US-09-869-708-276   |       | Sequence 14459, App |
| 13         | 88    | 10.6  | 2000 | 10     | US-09-919-497-10    |       | Sequence 276, App   |
| 14         | 88    | 10.6  | 2000 | 10     | US-09-880-107-3774  |       | Sequence 10, App    |
| 15         | 87    | 10.4  | 1209 | 9      | US-09-902-941-1918  |       | Sequence 3774, App  |
| 16         | 87    | 10.4  | 1209 | 9      | US-09-849-626-1918  |       | Sequence 1918, App  |
| 17         | 87    | 10.4  | 1209 | 9      | US-10-017-754-1918  |       | Sequence 1918, App  |
| 18         | 87    | 10.4  | 1238 | 9      | US-09-736-457-803   |       | Sequence 803, App   |
| 19         | 87    | 10.4  | 1238 | 9      | US-09-902-941-803   |       | Sequence 803, App   |

|    |     |                                                                |      |
|----|-----|----------------------------------------------------------------|------|
| QY | 263 | GTACATGCAGGTACCTGTGGCTGACTCCCTAACTCACGTCCTGTGACTCTTTTGACC      | 322  |
| Db | 456 | GTACATGCAGGTACCTGTGGCTGACTCCCTAACTCACGTCCTGTGACTCTTTTGACC      | 515  |
| QY | 323 | TATTGCTGACCATATCCACAGCGTGAGATGAAGCAGGGCCGTACTTTGCTGCACGTGC     | 382  |
| Db | 516 | TATTGCTGACCATATCCACAGCGTGAGATGAAGCAGGGCCGTACTTTGCTGCACGTGC     | 575  |
| QY | 383 | TGCTGGTGTGAGCGCTCAGCTGCCCTGTGCTCGCCTACCTCATGAAGTACCACGCCAT     | 442  |
| Db | 576 | TGCTGGTGTGAGCGCTCAGCTGCCCTGTGCTCGCCTACCTCATGAAGTACCACGCCAT     | 635  |
| QY | 443 | GTCCCTGCTGGAGCCGCACACGCTGGACCAAGTCATCGCGCCCATCATCCGACCCAAACG   | 502  |
| Db | 636 | GTCCCTGCTGGAGCCGCACACGCTGGACCAAGTCATCGCGCCCATCATCCGACCCAAACG   | 595  |
| QY | 503 | CGGCTTTTGGGAGCAGCTCATCCACATATGAGTTCCAAATTTGTTGGCAAGAACACTGTGCA | 562  |
| Db | 696 | CGGCTTTTGGGAGCAGCTCATCCACATATGAGTTCCAAATTTGTTGGCAAGAACACTGTGCA | 755  |
| QY | 563 | CATGGTCAGTTCCCGAGTGGGAATGATCCCTGCACATCTATGAGAAGGAAGTCGGTTTGTAT | 622  |
| Db | 756 | CATGGTCAGTTCCCGAGTGGGAATGATCCCTGCACATCTATGAGAAGGAAGTCGGTTTGTAT | 815  |
| QY | 623 | GATTCCTACTGTGAGCCATCCACGAGCCCTGCTATTTGGAGTCAGAGGTACAGATCTATTG  | 682  |
| Db | 816 | GATTCCTACTGTGAGCCATCCACGAGCCCTGCTATTTGGAGTCAGAGGTACAGATCTATTG  | 875  |
| QY | 683 | TTGATCTTACACCAAGATCCAACTTGAACATTCCTACTTTGTTGATACAGAAAAAACA     | 742  |
| Db | 876 | TTGATCTTACACCAAGATCCAACTTGAACATTCCTACTTTGTTGATACAGAAAAAACA     | 935  |
| QY | 743 | GATGATCCCTTTTATGAGCACAAAAAGAGTTGCTGTAGCTTTTAACCTTTATAATCCATT   | 802  |
| Db | 936 | GATGATCCCTTTTATGAGCACAAAAAGAGTTGCTGTAGCTTTTAACCTTTATAATCCATT   | 995  |
| QY | 803 | TTTTTTCAGATTAACATAATTGTGAGATGGTG                               | 834  |
| Db | 996 | TTTTTTCAGATTAACATAATTGTGAGATGGTG                               | 1027 |

## RESULT 2

```

US-09-764-869-257
; Sequence 257, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869.
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 257
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1104)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-869-257

```

|    | Query Match           | 96.6%                                                          | Score 805.6;        | DB 10;        | Length 1140;   |
|----|-----------------------|----------------------------------------------------------------|---------------------|---------------|----------------|
|    | Best Local Similarity | 99.5%;                                                         | Pred. No. 1.1e-244; |               |                |
|    | Matches 808;          | Conservative                                                   | 0;                  | Mismatches 4; | Indels 0; Gaps |
| Qy | 23                    | CATTCTTTGGTGTTCGTGACTGCTGACCACTGACCCACCGCGCTTGATGACAGCACCCCTC  | 82                  |               |                |
|    |                       |                                                                |                     |               |                |
|    |                       |                                                                |                     |               |                |
| Db | 216                   | CAGCGCTTCGTGCTTCGTGACTGCTGACCACTGACCCACCGCGCTTGATGACAGCACCCCTC | 275                 |               |                |

|    |     |                                                                |      |
|----|-----|----------------------------------------------------------------|------|
| Qy | 83  | GTGTGCGCTTCCAGTTCAGTTTCGGGACGCCCTCAGTTCAGGGCGCTCTCGCAGATACCAA  | 142  |
| Db | 276 | GTGTGCGCTTCCAGTTCAGTTTCGGGACGCCCTCAGTTCAGGGCGCTCTCGCAGATACCAA  | 335  |
| Qy | 143 | AAGCCTCTATATCAGCAATGCTGTGGCGCCCAACAACAGCTCATGCTGTCTAGCAACCA    | 202  |
| Db | 336 | AAGCCTCTATATCAGCAATGCTGTGGCGCCCAACAACAGCTCATGCTGTCTAGCAACCA    | 395  |
| Qy | 203 | GATCACCATGGTCATCAATGCTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA  | 262  |
| Db | 396 | GATCACCATGGTCATCAATGCTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA  | 455  |
| Qy | 263 | GTACATGAGGTACCTGTGGCTGACTCCCTTACTCACTCACTGCTCTGTGACTCTTTTGACCC | 322  |
| Db | 456 | GTACATGAGGTACCTGTGGCTGACTCCCTTACTCACTCACTGCTCTGTGACTCTTTTGACCC | 515  |
| Qy | 323 | TATTTCGTGACCATATCACAGCTGGAGATGAGCAGGGCGGTACATTTGCTGCACGTGC     | 382  |
| Db | 516 | TATTTCGTGACCATATCACAGCTGGAGATGAGCAGGGCGGTACATTTGCTGCACGTGC     | 575  |
| Qy | 383 | TGCTGGTGTGAGCGCTCAGCTGCCCTGTGCCYGGCTACCTCATGAAGTACCACGCCAT     | 442  |
| Db | 576 | TGCTGGTGTGAGCGCTCAGCTGCCCTGTGCCYGGCTACCTCATGAAGTACCACGCCAT     | 635  |
| Qy | 443 | GTCCCTGTGGAGCCGCACACGTGGACCAAGTCATCGCGGCCCATCATCCGACCACAAG     | 502  |
| Db | 636 | GTCCCTGTGGAGCCGCACACGTGGACCAAGTCATCGCGGCCCATCATCCGACCACAAG     | 695  |
| Qy | 503 | CGGCTTTGGGAGCAGCTCATCACATGATGAGTTCCAAATGTTTGGCAAGAACACTGTGCA   | 562  |
| Db | 696 | CGGCTTTGGGAGCAGCTCATCACATGATGAGTTCCAAATGTTTGGCAAGAACACTGTGCA   | 755  |
| Qy | 563 | CATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCGGTTTGAT   | 622  |
| Db | 756 | CATGGTCAGTTCCCGAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCGGTTTGAT   | 815  |
| Qy | 623 | GATTCCACTGTGAGCCATCCACAGAGCCCTGTCATTTGGAGTCAGAGGTACAGATCTATTG  | 682  |
| Db | 816 | GATTCCACTGTGAGCCATCCACAGAGCCCTGTCATTTGGAGTCAGAGGTACAGATCTATTG  | 875  |
| Qy | 683 | TTGATCTTACACCAAGATCCAACTTGAACATTTCTACTTTTGTGTGATACAGAAAAACA    | 742  |
| Db | 876 | TTGATCTTACACCAAGATCCAACTTGAACATTTCTACTTTTGTGTGATACAGAAAAACA    | 935  |
| Qy | 743 | GATGATGCCTTTTATGAGCACAAAAAGAGTTGCTGTAGCTTTTAACTTTTAAATCCATT    | 802  |
| Db | 936 | GATGATGCCTTTTATGAGCACAAAAAGAGTTGCTGTAGCTTTTAACTTTTAAATCCATT    | 995  |
| Qy | 803 | TTTTTTCAGATTAACTAAATTGTGAGATGGTG                               | 834  |
| Db | 996 | TTTTTTCAGATTAACTAAATTGTGAGATGGTG                               | 1027 |

### RESULT 3

```

US-10-091-504-1735
; Sequence 1735, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1735
; LENGTH: 7537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1735

```

Query Match 96.48; Score 804; DB 9; Length 7537;

```

; SEQ ID NO 1735
; LENGTH: 7537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1735

Query Match          96.4%; Score 804; DB 10; Length 7537;
Best Local Similarity 99.4%; Pred. No. 1.4e-243;
Matches 807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY   23 CATTCCTGGTGTTCGCTGACTGCTGCCAGTACACCAGCCACCGCCTTGATGACAGCACCCCTC 82
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   3810 CAGCCTTCGGTGTTCGCTGAGCTGCTGACCAGTACCACCGCCTTGATGACAGCACCCCTC 3869

QY   83 GTGTGCCTTCCCAGTTTCAGATTCCGGCAGCCCTCAGTTCAGGGGCCCTCTCGCAGATAACCAA 142
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   3870 GTGTGCCTTCCCAGTTTCAGTTCCGCGACCCCTCACTCAGGGGCCCTCTCGCAGATAACCAA 3929

QY   143 AAGCCTGTATATACGAATGGTGTGGCGGCCAACAAACAAGCTCATGTGCTTAGCAACCA 202
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   3930 AAGCCTGTATATACGAATGGTGTGGCGGCCAACAAACAAGCTCATGTGCTTAGCAACCA 3989

QY   203 GATCACCATGGTCAATCATGTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 262
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   3990 GATCACCATGGTCAATCATGTCTCAGTGGAGGTAGTGAACACCTTGTATGAGGATATCCA 4049

QY   263 GTACATCAGGTACCTGTGGCTGACTCCCCTAACCTACAGTCTCTGTGACTTCTTTTGACCC 322
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4050 GTACATCAGGTACCTGTGGCTGACTCCCCTAACCTACAGTCTCTGTGACTTCTTTTGACCC 4109

QY   323 TATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCGCGTACTTTGCTGCACTGTGC 382
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4110 TATTGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCGCGTACTTTGCTGCACTGTGC 4169

QY   383 TGCTGGTGTGAGCGGCTCAGCTGCCCTGTGCCTCGCCCTACCTCATAGTACACAGCCAT 442
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4170 TGCTGGTGTGAGCGGCTCAGCTGCCCTGTGCCTCGCCCTACCTCATAGTACACAGCCAT 4229

QY   443 GTCCTCTGCTGAGCGCCACACAGTGCAGCAAGTCAATCGCGGCCCATCATCCGACCCCAACAG 502
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4230 GTCCTCTGCTGAGCGCCACACAGTGCAGCAAGTCAATCGCGGCCCATCATCCGACCCCAACAG 4289

QY   503 CGGCTTTTGGGAGAGCTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACAAGTGTGCA 562
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4290 CGGCTTTTGGGAGAGCTCATCCACTATGAGTTCCTCAATTTGTTGGCAAGAACAAGTGTGCA 4349

QY   563 CATGGTCAGTTCCTCCAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 622
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4350 CATGGTCAGTTCCTCCAGTGGGAATGATCCCTGACATCTATGAGAAGGAAGTCCGTTTGAT 4409

QY   623 GATPCCACTGTGAGCCATCCCACGAGCCCTTGCATTTGGAGTCAGAGGTACAGATCTATTG 682
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4410 GATPCCACTGTGAGCCATCCCACGAGCCCTTGCATTTGGAGTCAGAGGTACAGATCTATTG 4469

QY   683 TTGATCTTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGTATACAGAAAAAACA 742
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB   4470 TTGATCTTTACACCAAGATCCAAACTTGAACATTTCTACTTTTGTGTATACAGAAAAAACA 4529

```

```
Db   4530 GATGATGCCCTTTTATGAGCACAAAAAAGAGTGCTGTAGCTTTTTTAACCTTTATAATCCATT    4589
      |||||
Qy   803 TTTTTTCAGATTAAACTAAATTGCAGATGGNG    834
      |||||
Db   4590 TTTTTTAAGATTAAACTAAATTGCAGATGGNG    4621
      |||||
```

RESULT 5  
US-09-864-761-26715/c  
; Sequence 26715, Application US/09864761  
; Patent No. US20020048763Al  
; GENERAL INFORMATION:  
; APPLICANT: Penn. Shalron G.



```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

```

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

ORGANISM: Homo sapiens

## RESULT 7

US-10-044-205A-41  
; Sequence 41, Application US/10044205A  
; Patent No. US20020123464A1  
; GENERAL INFORMATION:  
; APPLICANT: KAPELLEL-LTBERMANN, Rosana



```

; NAME/KEY: misc feature
; LOCATION: (191)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (221)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-8474

Query Match 17.9%; Score 149.6; DB 10; Length 239;
Best Local Similarity 91.3%; Pred. No. 3.2e-37;
Matches 190; Conservative 0; Mismatches 13; Indels 5; Gaps

Qy 617 TTTGATGATTCCACTGTGAGCCATCCACGAGGCCCTGCATTTGGAGTCAGAGGTACAGAT 676
Db 32 TTTAATGATTCCACTGTGAGCCATCCACGAGGCCCTGCATTTGGAGTCAGAGGTACAGAT 91
Qy 677 CTATTGTTGATCTTACACCAAGATCAAACTTCGAACATTCCTACTTTGTTGTATACAGAAA 736
Db 92 CTATTGTTGATCTTACACCAAGATCCAACTTGAACATTNTACTTTATTGGTACAGAAA 151
Qy 737 AAAACAGATGATGCCTTTT-ATGAGCACAAAAAAGAGTT-GCTGTAGCTTTT---AACTT 791
Db 152 AAAACAGATGATGCCTTTAAATGAGCACAAAAAAGNGTNGCTGTAGCTTGTGACTTT 211
Qy 792 TATAATCCATTTTTTTTCAGATTAAC 819
Db 212 AATAATCCATTTTTTTTCAGGTTAAACT 239

RESULT 11
US-09-864-761-14459
; Sequence 14459, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

```

```

> APPLICANT: Hanzel, David K.
> APPLICANT: Chen, Wensheng
> TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
> TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
> FILE REFERENCE: Aeomica-X-1
> CURRENT APPLICATION NUMBER: US/09/864,761
> CURRENT FILING DATE: 2001-05-23
> PRIOR APPLICATION NUMBER: US 60/180,312
> PRIOR FILING DATE: 2000-02-04
> PRIOR APPLICATION NUMBER: US 60/207,456
> PRIOR FILING DATE: 2000-05-26
> PRIOR APPLICATION NUMBER: US 09/632,366
> PRIOR FILING DATE: 2000-08-03
> PRIOR APPLICATION NUMBER: GB 24263.6
> PRIOR FILING DATE: 2000-10-04
> PRIOR APPLICATION NUMBER: US 60/236,359
> PRIOR FILING DATE: 2000-09-27
> PRIOR APPLICATION NUMBER: PCT/US01/00666
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00667
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00664
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00669
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00665
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00668
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00663
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00662
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00661
> PRIOR FILING DATE: 2001-01-30
> PRIOR APPLICATION NUMBER: PCT/US01/00670
> PRIOR FILING DATE: 2001-01-30

```



```

: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine
: SEQ ID NO 14459

```

|                       |              |                    |                 |             |
|-----------------------|--------------|--------------------|-----------------|-------------|
| Query Match           | 11.8%;       | Score 98.4;        | DB 10;          | Length 467; |
| Best Local Similarity | 57.0%;       | Pred. No. 8.6e-21; |                 |             |
| Matches 180;          | Conservative | 0;                 | Mismatches 136; | Indels 0;   |
|                       |              |                    | Gaps            | 0;          |

```

RESULT 12
US-09-969-708-276
; Sequence 276, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 558
; SOFTWARE: patentIn version 3.0
; SEQ ID NO 276
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-276

```

|                           |        |                    |           |              |
|---------------------------|--------|--------------------|-----------|--------------|
| Query Match               | 10.6%  | Score 88;          | DB 10;    | Length 2000; |
| Best Local Similarity     | 51.0%; | Pred. No. 4.6e-17; |           |              |
| Matches 208; Conservative | 0;     | Mismatches 200;    | Indels 0; | Gaps 0;      |

  

|    |      |                                                                   |      |
|----|------|-------------------------------------------------------------------|------|
| Qy | 147  | CTGTATATACAGCAATGCTGTGGCGGCCAACAAAGCTCATGCTGTCTAGCAACCAGATC       | 206  |
| Db | 774  | CTGTACTGGGCAGTGGCTATACGCTTTCCCGAAGCACATGCTGGATGCTTGGGCGNTA        | 833  |
|    |      |                                                                   |      |
| Qy | 207  | ACCATGTTCATCAATGTCTCAGTGAGGTAGTGAACACCTTGTATGAGGATATCCAGTAC       | 266  |
| Db | 834  | ACTGCCTTGATCAACGCTCTCAGCCAATTGTCCCAACCACTTTTGAGGGTCACTACCAAGTAC   | 893  |
|    |      |                                                                   |      |
| Qy | 267  | ATGCAGGTACTCTGGCTGTGACTCCCCTTAACCTACGCTCTCTGTGACTTCTTTTGACCCCTATT | 326  |
| Db | 894  | AAGAGCATCCCTGTGTGAGGACAACCAAGCAGACATAGCTCCTGGTTGGTCAACGAGGCC      | 953  |
|    |      |                                                                   |      |
| Qy | 327  | GCTGACCATAITCCACACGCTGGAGATGAAGCAGGGCCGTACTTTGTCTGCACCTGTGCTGCT   | 386  |
| Db | 954  | ATTGACTTCATAGACTCCATCAAGAATGCTGGAGGAAGGGTGTTTGTGCCACTGCCAGGCA     | 1013 |
|    |      |                                                                   |      |
| Qy | 387  | GGTGTGAGCGCTCAGCTGCGCTGTGCGCTGCCCTACCTCATGAAGTACCAGGCCNATGCC      | 446  |
| Db | 1014 | GGCAATTTCCGGTCAGCACCACTGCTGCTTGTACCTTATGAGGACTAATCGAGTCAAG        | 1073 |
|    |      |                                                                   |      |
| Qy | 447  | CTGTGAGCGCCACACGCTGGACCAACTCATGCGCGGCCCATATPCCGACCAACACAGCGGC     | 506  |
| Db | 1074 | CTGACAGGCGCTTTGAGTTTGTGAACGAGGCGAAGCATCATCTCCCAACTTCAGC           | 1133 |
|    |      |                                                                   |      |
| Qy | 507  | TTTTGGGAGCAGCTATCCCACTATGAGTTCOAATTTGTTGGCAAGAC                   | 554  |
| Db | 1134 | TTCATGGCCAGCTGCTGTCAGTTTGAGTCCCAAGTGTGGCTCCGCAAC                  | 1181 |

```

RESULT 13
US-09-919-497-10
; Sequence 10, Application US/09919497
; Patent No. US2002010662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-497-10

```

|    | Query Match           | 10.6%                                                         | Score 88;          | DB 10;          | Length 2000;      |
|----|-----------------------|---------------------------------------------------------------|--------------------|-----------------|-------------------|
|    | Best Local Similarity | 51.0%;                                                        | Pred. No. 4.6e-17; |                 |                   |
|    | Matches 208;          | Conservative                                                  | 0;                 | Mismatches 200; | Indels 0; Gaps 0; |
| Qy | 147                   | CTGTATATCAGCAATGCTGGCGCCGCAACAACAGCTCATGCTGTCTAGCAACACGATC    | 206                |                 |                   |
|    |                       |                                                               |                    |                 |                   |
|    |                       |                                                               |                    |                 |                   |
| Db | 774                   | CTGTACTGGCGATGGGTATCAGCTTCCGGCAGACATGCTGGATGCTTGGGCGATA       | 833                |                 |                   |
| Qy | 207                   | ACCATGTCATCAATGCTCTCAGTGGAGGTAGTGAAACACCTTGTATGAGGATATCCAGTAC | 266                |                 |                   |
|    |                       |                                                               |                    |                 |                   |
|    |                       |                                                               |                    |                 |                   |
| Db | 834                   | ACTGCTTGATCAACGCTCTCAGCCAATTGTGCCAACCATTTTGAGGGTCACTACCAGTAC  | 893                |                 |                   |
| Qy | 267                   | ATGCAGGTACTCTGGGCTGACTCCCTTAATCAGTCTCTGTGACTTCTTTGACCCCTATT   | 326                |                 |                   |
|    |                       |                                                               |                    |                 |                   |
|    |                       |                                                               |                    |                 |                   |
| Db | 894                   | AAGAGATCCCTGTGGAGGACAACCAAGCAGACATCAGTCTCCTGGTTCAACAGGGCC     | 953                |                 |                   |
| Qy | 327                   | GCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTACTTTGCTGCACCTGTGCTGCT | 386                |                 |                   |
|    |                       |                                                               |                    |                 |                   |
| Db | 954                   | ATTCACTTCATAGACTCCATCAAGATGCTGGAGGAGGGTGTGTGCTCACTGCCAGGCA    | 1013               |                 |                   |



```

RESULT 15
US-09-902-941-1918
: Sequence 1918, Application US/09902941
: Patent No. US20020172952A1
: GENERAL INFORMATION:
: APPLICANT: Henderson, Robert A.
: APPLICANT: Wang, Tongtong
: APPLICANT: Watanabe, Yoshihiro
: APPLICANT: Johnson, Jeffrey C.
: APPLICANT: Retter, Marc W.
: APPLICANT: Warnerakis, Margarita
: APPLICANT: Carter, Darrick
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: Bangur, Chaitanya S.
: APPLICANT: McNabb, Andria
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.478C17
: CURRENT APPLICATION NUMBER: US/09/902,941
: CURRENT FILING DATE: 2001-07-10
: NUMBER OF SEQ ID NOS: 2002
: SOFTWARE: FastSeq for Windows version 4.0
: SEQ ID NO 1918
: LENGTH: 1209
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-902-941-1918

Query Match 10.4%; Score 87; DB 9; Length 1209;
Best Local Similarity 49.9%; Pred. No. 6.8e-17;
Matches 219; Conservative 0; Mismatches 220; Indels 0; Gaps

Qy 146 CCTGTATATACGAAATGGTGTGCCGCCAACACAAAGCTCATGCTGTCTAGCAACACGAT 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 627 CCTCTACTCGGAGTGCCTTACCATTGCTGCCGGAGAGACATGCTGGACGCCCTGGGCAT 686
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 206 CACCATGGTCATCAANTCTCTCAGTGAGGTAGTGAACACCTTGTATGAGGATATCCAGTA 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 687 CACGGCTCTCTTGAATGTCTCTCGGACTGCCCCAACCAACCATTTGAAGGACATATCAGTA 746
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 266 CATCAGGTACCTGTGGCTGACTCCCTCAACTCAGCTCTCTGTGACTTCTTTTGACCCCTAT 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 747 CAAGTGCATCCAGTGAAGATAAACACAAAGCCGACATCAGCTCCTGTTTGGGAAGC 806
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 326 TGCTGACCATATCCACAGCGTGGAGATGAAGCAGGGCCGTACTTTGCTGCACTGTGCTGC 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 807 CATAGAGTACATCGATCCCTGAAGACTGCGCGTGGCGCGTGCCTGCTGCACTGCCAGGC 866
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 386 TGGTGTGAGCGCTCAGCTGCCCTGTGCTCCCTTACCTCATGAAGTACCAGCCCATGTC 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 867 GGGCATCTCGCGGTGCGCCACCATCTGCTCTGCCCTACCTGATGATGAAGAACGGGTGAG 926
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 446 CCTGCTGGAGCGCCACACGCTGGACCAAGTCAATCGCGGCCCATCATCCGACCCCAACAGCGG 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 927 GCTGGAGAGCGCTTCGAGTTCCTTAAGCAGCGCCGACGACATCATCTGCCCAACTTCAG 986
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 506 CTTTGGGAGCAGCTCATCCACTATGAGTTCCAATTTGTTGGCAAGAACACTGTGCACAT 565
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 987 CTTTCATGGGCAGCTGCTGCAGTTCGAGTCCAGGTCCTCCAGGTGCTGGCCAGCTCTGTGTCGCGGA 1046
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 566 GGTCAAGTTCCTCCAGTGGGA 584
   ||| ||| ||| ||| |||
Db 1047 GGCTGTAGCCCTCGGGA 1065
   ||| ||| ||| ||| |||

```

Search completed: April 21, 2003, 18:31:47  
Job time : 104 secs

**THIS PAGE BLANK (USPTO)**